Louisiana State University
LSU Digital Commons

LSU Doctoral Dissertations | Graduate School

2012

Defense gland protein from Coptotermes formosanus

Ioan-Horia Negulescu
Louisiana State University and Agricultural and Mechanical College, inegul1@tigers.lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Recommended Citation
Negulescu, Ioan-Horia, "Defense gland protein from Coptotermes formosanus" (2012). *LSU Doctoral Dissertations*. 3509.
https://digitalcommons.lsu.edu/gradschool_dissertations/3509

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contactgradetd@lsu.edu.

DEFENSE GLAND PROTEIN FROM *COPTOTERMES FORMOSANUS*

A dissertation
Submitted to the Graduate Faculty of the
Louisiana State University
Agricultural and Mechanical College
In partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy

In

The Department of Biological Sciences

By
Ioan Horia Negulescu
B.A., Case Western Reserve University, 1994
May, 2012

DEDICATION

This work is dedicated to my beloved wife Catalina and my children Andra and Alexandru without whom I would not have been who I am.

ACKNOWLEDGEMENTS

I would like to thank first and foremost to Dr. Roger Laine who believed in me, accepted to be my professor and mentor, and made me part of his distinguished scientific career. I wish to thank my research committee members Dr. Gregg Henderson for his infinite patience and help; Dr. Evanna Gleason for her moral boosting and understanding, accepting to make me part of her lab meetings; Dr. Yong-Hwan Lee who is helping me with the X-ray diffraction studies and provided background for the NMR studies.

A special thank you to Dr. Megan McNaughton, who arrived at LSU just when I needed her most and was always available for discussions, and Thomas Garner who performed the NMR experiments and helped me understand the results.

Lucas Veillon, thank you for your company and being always ready to help.

Last but not least, I thank Dr. Betty Zhu for everything she has done for me. Her retirement left a void in the lab and in my heart.

TABLE OF CONTENTS

ABSTRACT

Coptotermes formosanus Shiraki is the most invasive termite species and has the largest economic impact of all the subterranean termites in the US. Although a lot of work has been done on the determination of chemical composition of frontal gland secretion from different families of termites, the protein component of those secretions was relatively overlooked.

This study started with the observation that proteins are present in the frontal gland secretion of *Coptotermes formosanus* Shiraki soldiers. The goal was to characterize one of these proteins by identifying its DNA sequence which would lead to protein sequence, function, structure and role it plays in the defense of *Coptotermes formosanus* Shiraki.

The protein cloned named TFP4 and analysis of its sequence revealed a molecular weight of 6853 Daltons. The sequence was compared with those of known proteins in NCBI's data base and similarities with other proteins analyzed. Two functions were assigned to TFP4, lysozyme and serine protease inhibitor.

The lysozyme function was assayed using a zymogram and a fluorometric assay. The zymogram showed clear zones where TFP4 had migrated indicating the β-(1,4) glucosidic linkages between N-acetyl-muramic acid and N-acetyl-D-glucosamine in the cell wall of *Micrococcus lysodeikticus* were hydrolysed. The fluorometric assay showed the fluorescent molecule which was embedded in the cell wall of *Micrococcus lysodeikticus* was released in solution upon addition of TFP4.

The serine protease assays tested the effect TFP 4 addition has on the reaction rates of trypsin, chymotrypsin and elastase with their respective substrates. The rate was unchanged in trypsin's assay but chymotrypsin and elastase were inhibited by TFP4.

In conclusion, we strongly suggest the existence of a protein of 6853 KDa with dual functions, lysozyme and serine protease inhibitor (chymotrypsin and elastase) in the defensive secretion of frontal gland of *Coptotermes formosanus* Shiraki soldiers. The inhibitor function is assigned to an exposed loop from Cys4 to Cys16 which functions as a substrate analog and employs a competitive inhibition mechanism for the inhibition of chymotrypsin while the lysozyme function is putatively assigned to a pocket between the β-hairpin structure containing Gln23 and the region at the C-terminus containing Glu45.

CHAPTER ONE: LITERATURE REVIEW

Termites are a group of eusocial insects which are classified under the cockroach order Blattodea, in the epifamily Termitoidae. Over 2800 species of termites are recognized and are classified in seven families (Engel and Krishna, 2004).

Strong morphological and DNA evidence suggests the termites are closely related to cockroaches having evolved from a primitive cockroach ancestor approximately 200 million years ago (Krishna, 1990) and are, in fact, highly modified, eusocial, wood eating cockroaches (Weesner *et al.,* 1960; Lo *et al.,* 2003; Ware *et al.,* 2008). These soft bodied, mostly blind insects use chemical communication for caste differentiation and recognition, nest building, food finding and defense (Stuart *et al.,* 1970; Moore *et al.,* 1974).

As social insects, like ants, bees and wasps, there is a strict hierarchal system, a division of labor generated by the formation of castes in a colony. In general, castes contain a king and a queen (reproductives), nymphs (semi-mature young), workers and soldiers (Wilson, 1971). All individuals have the genetic information for all caste types in their genome so they have the potential to become members of any caste through caste differentiation, a phenomenon known as polyphenism (Noirot, 1991).

Polyphenisms are alternative morphological phenotypes that differentiate in response to environmental conditions. Larvae differentiate into either workers or nymphs; workers can undergo worker to worker molts, differentiate into presoldiers or differentiate into ergatoid neotenics. Nymphs can regress into workers, differentiate into fully winged reproductives or differentiate into brachypterous neotenics. Full winged, eyed alates are primary reproductives, winged and eyeless brachypterous neotenics are secondary

reproductives while wingless, eyeless ergatoid neothenics are third form reproductives. Both extrinsic and intrinsic factors influence a colony's differentiation into castes (Sharf *et al.,* 2010).

The extrinsic factors that were recently confirmed as influencing a colony's structure include temperature, worker/soldier proportion and food quantity. Higher temperatures over 27^0C induce differentiation possibly by influencing Juvenile Hormone (JH) biosynthesis. High food quality is associated with high JH titers but food abundance is associated with low JH titers (Lo *et al.,* 2009).

Intrinsic factors include the Juvenile Hormone, secreted by the corpus allatum. The JH titers peak just before molting of workers into presoldiers and it typically happens just before swarming. The latest research in the termite's caste control looks at various factors that influence JH action, both exogenous and endogenous. The morphogenetic JH is a central player in the population structure control and it has been shown to play a role in the control of development and metamorphosis, diapauses, migratory behavior, wing length, seasonal development, reproduction (Mao *et al.,* 2005; Ishikawa *et al.,* 2007.; Vargo *et al.,* 2009).

While the actions of JH were described before, the hexamerin proteins were newly identified as controlling the levels of JH. Hex-1 can bind JH and Hex-2 has receptor-like qualities and a unique hydrophilic insertion with unknown function. Hexamerins were previously described as nutrient storage proteins. Indeed, when feeding is at its highest no presoldier differentiation is observed, JH titers being low. Silencing of both hexamerins with RNAi, without exogenous JH, leads to a significant increase in JH dependent presoldier differentiation (Sharf., 2010). Thus, hexamerins play a role in the

social regulation of a colony that is regulated by nutrition, endogenous factors (JH titers) and environment (ambient temperature).

The soldier caste contains highly specialized individuals developed specifically for colony defense against predators but the soldiers are incapable of performing other tasks and they have to be fed by workers (Prestwich, 1978). They are best described as "a caste of nightmarish monsters, which recall the most fantastic revelries of Hieronymous Bosch, Breughel the Elder and Callot... Its own weapons have not been borrowed, like ours, from the external world; it has done better . . . . It has created those weapons out of its own body " (Maeterlinck, 1939).

Termites have enemies among both invertebrates and vertebrates. The most studied and most important termite predators are the ants (Hill, 1922; Blake, 1941; Weber, 1964; Sheppe, 1970), but other invertebrate enemies include platyhelmintes (Sheppe, 1970), members of Arachnida - scorpions (Dracott, 1919), spiders (Sheppe, 1970), myriapodes (Sheppe, 1970), other insects from the orders of Diptera (Carpenter, 1919), Coleoptera (Kalshoven, 1955) and Hymenoptera (Guido and Ruffinellli, 1958). Vertebrate predators include amphibians (Dimmit, 1975), reptiles (Hoogmoed, 1973), birds (Rowan, 1971), marsupials (Calaby, 1960) and insectivores (Kingdom, 1971). Defense mechanisms employed by termites against predators are both non-chemical or mechanical (Deligne, 1971) and chemical (Quennedey, 1975).

Mechanical defense takes advantage of evolved, specialized mandibles and can be achieved by slashing, crushing, piercing or snapping (Prestwich, 1984). Other tactics are phragmosis, the physical plugging of entry holes to colony by soldiers literally using their heads, autothysis, the forcing of labial gland secretion through a weakness in the

abdominal wall (Deligne *et al.,* 1982) and abdominal dehiscence, the split of abdominal wall (Mill, 1982).

There are three fundamental mechanisms for chemical defense. In the first mechanism the soldier bites or pierces the attacker using its mandibles while, at the same time, it injects or applies a chemical secretion from a frontal gland. A second technique is the application of a large quantity of topically active poison secreted by the labial gland and "daubed" on the attacker using an extended labrum with broad hairs which acts like a brush. Finally, a third mechanism, employed by Nasutitermes, is to eject a viscous solution that contains irritants and glue which immobilizes the enemy (Quennedey, 1973; Prestwich, 1979). The secretions used in defense are produced by the salivary (labial) glands, the frontal gland and the labral (cibarial) gland.

The salivary (labial) glands used in defense are a pair of hypertrophied glands with reservoirs extending into the abdomen (Quennedey, 1981). The secretion is rich in quinones and proteinaceous material and it often entangles the predator because it hardens after air exposure (Wood, 1975).

The frontal gland is well developed in Rhinotermitidae and Termitidae and is an unpaired cephalic organ which opens through a frontal pore or fontanelle (Deligne, 1984). It reaches deep into the abdominal cavity pushing other organs and the secretion may weigh as much as one third of the total body weight (Waller and La Fage, 1987).

The labral gland opening overlaps the mandible and its secretion is used to impregnate the mandibular blades thus getting access to the wound during a fight (Quennedey, 1983).

The chemicals present in these secretions come from three main pathways (figure 1.1) - mevalonate (terpenoid), fatty acid metabolism and amino acid or carbohydrate metabolism. (Prestwich, 1984).

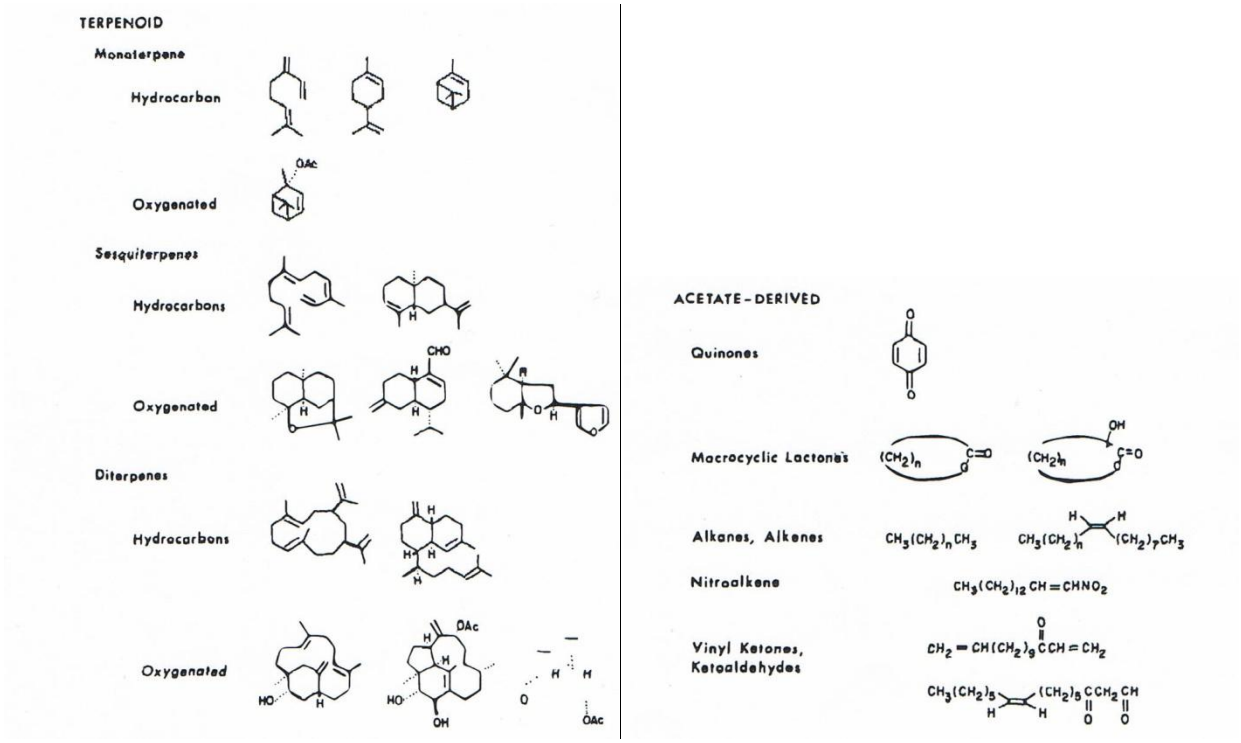
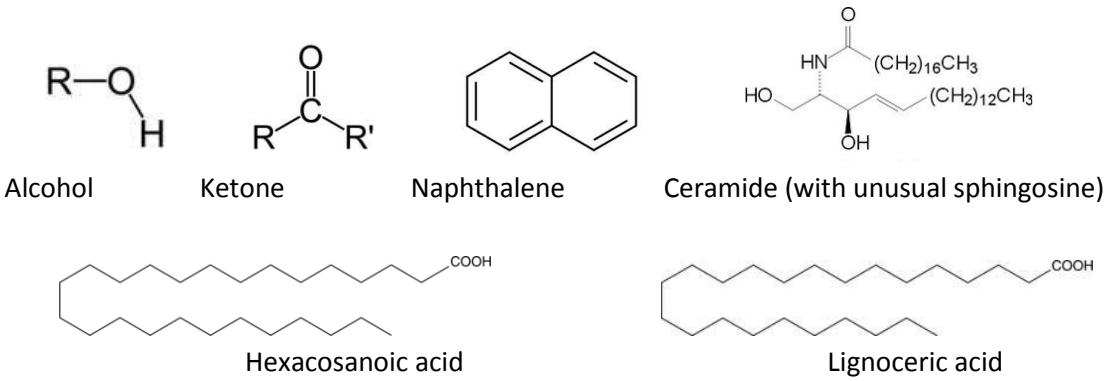


Figure 1.1 Representative chemicals from termite defense gland secretion (from Prestwich, 1984)

Termites are invasive species and their diet consists of wood fiber (from live trees, wood structures), crops, plants and other materials which contain cellulose. Protozoa, bacteria and yeasts living in the digestive system digest cellulose and provide energy and nutrition (Cabrera *et al.,* 2001). The economic impact of termite activity reaches up to $11 billion in losses annually in the United States (Su *et al.,* 2002). The most destructive and economically important genera are Coptotermes and Reticulitermes of the subterranean termite family of Rhinotermitidae. Members of Coptotermes subfamily are very active in the Southern States. Coptotermes formosanus Shiraki, the Formosan subterranean termite, was introduced in the US in late 1940s-early 1950s when it was

brought in military cargo and ballast by ships from Asia following World War II (La Fage, 1985). New Orleans, La. was one of the most active ports at the time and C. formosanus benefited from a warm and humid climate to establish colonies in structures and urban trees (Osbrink, 1999)

In a *C.formosanus* colony soldiers can comprise up to 60% of its members (Haverty, 1977).They employ both mechanical and chemical defense mechanisms. Their long falcate mandibles are modified for crushing and, during combat, the milky defense secretion of its frontal gland is ejected through a frontal pore onto the attacker, rapidly stiffening in the air (Spaton and Prestwich, 1981).



Alcohol Ketone Naphthalene Ceramide (with unusual sphingosine)

Hexacosanoic acid Lignoceric acid

Figure 1.2 Components of frontal gland secretion of C. formosanus

Various components of frontal gland secretion were identified by our laboratory through the years (figure 1.2): hexacosanoic acid and lignoceric acid (Chen *et al.,* 1999), hydrocarbons, aromatic compounds, ketones, alcohols (Zhang *et al.,* 2006), ceramides (Ohta *et al.,* 2007) and naphthalene (Chen, 1998).

Although a lot of work has been done on the determination of chemical composition of frontal glands from different families of termites, the protein component of secretions was relatively overlooked. Recent work has generated interest in caste-specific gene

expression. Miura et al. describe SOL1, a protein which is found only in the mandibular glands of *Hodotermopsis japonica,* but not in any other stages of differentiation into presoldiers (Miura *et al.,* 1999). Hojo et al. identified Ntsp 1, a secretory carrier protein present only in the epithelial cells of the frontal gland reservoir of *Nasutitermes takasagoensis* soldiers (Hojo *et al.,* 2005). In *Reticulotermes flavipes*, the most common termite in the United States, Scharf describes the presence of soldier-specific gene expression of transcription and translation factors with significant sequence homology to the *bicaudal* and *bric-a-brac* genes in Drosophila which function in embryonic pattern formation (Scharf *et al.,* 2003).

Coptotermes formosanus Shiraki is the most invasive termite species and has the largest economic impact of all the subterranean termites in the US. It is not a native species; it changes the ecological equilibrium by displacing native species and is classified as a pest. Various control methods and termiticides were tried and are in use but there is a need for the discovery of new, more efficient methods of termite control. The more we know about their biology, the more we can use the knowledge against them. Also, we can use the knowledge derived from these studies for new synthetic compounds by taking advantage of the hundreds of millions of years of evolution. Untold numbers of drugs were created using the information derived from naturally occurring biological compounds.

This study started with the observation that proteins are present in the termite frontal gland secretion of *Coptotermes formosanus* Shiraki soldiers. The goal of this study was to characterize one of these proteins by identifying its DNA sequence which would lead

to protein sequence, function, structure and role it plays in the defensive arsenal of *Coptotermes formosanus* Shiraki.

CHAPTER TWO: ISOLATION AND CLONING OF TFP4 (TERMITE FRONTAL GLAND PROTEIN 4)

Introduction

A protein gel of the defense secretion revealed the existence of 5-8 proteins. The N-terminal sequence determination of TFP4 was performed by Dr. Richard Cook at Baylor University Medical School in Houston, TX and was determined as EDC/WQLFC/WPMIY. The plan of research was to isolate total mRNA from termite soldiers and then isolate and sequence the cDNA corresponding to TFP4 protein.

The total mRNA isolated from the soldiers of C. formosanus Shiraki was used to create a cDNA library and further isolate the corresponding cDNA of interest using specific primers.

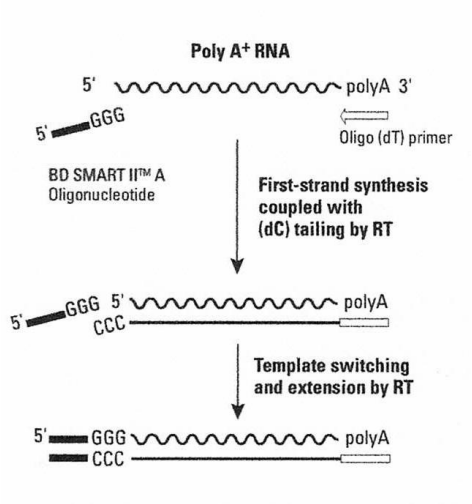


Figure 2.1 General strategy for converting total mRNA to cDNA library using Reverse Transcriptase. After reverse transcriptase reaches the end of the mRNA template it adds several dC residues. BDSMART II A oligonucleotides anneal to the dC residues and serve as extended templates (from BD Biosciences)

The BD SMART (Switching Mechanism At 5' end of RNA Transcript) RACE (Rapid Amplification of cDNA Ends) technique of cDNA synthesis from BD Biosciences

Clontech was used because the N-terminal amino acid sequence of the protein of interest was known.

The strategy used takes advantage of the poly-A tails of mRNAs to generate a cDNA library using a 3'-primer which is complimentary to the mRNAs' ends.

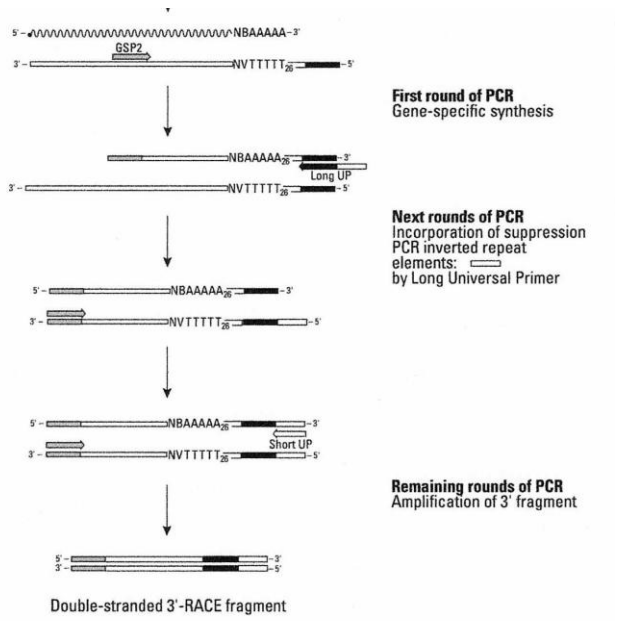


Figure 2.2 Gene specific cDNA synthesis using primers derived from the known N-terminal protein sequence of TFP4. Long UP primer is used as 3' primer to replace the poly-T primer used in the reverse transcription step and the 5' primer is specific to the sequence of interest (from BD Biosciences).

To isolate the cDNA of interest a PCR reaction was performed. The 5'-primer used was a mix of oligonucleotides complimentary to the 5' end, based on the N-terminal sequence of the target protein. The 3' primer was a mix of oligonucleotides compatible with the 3' end of cDNAs in the library. After the reaction was completed the PCR product was purified using the QIAquick PCR purification kit (Qiagen Sciences, Md.)

The TFP4 cDNA was then inserted into the pET46Ek/LIC vector (Novagen brand, EMD Biosciences) for expression of protein using isopropyl-1-thio-β-D-galactopyranoside

(IPTG). This vector was designed for ligation independent cloning (LIC) and is engineered to express the target protein immediately downstream of an enterokinase (Ek) cleavage site, enzyme used for removal of the cloned protein from the His-tag.

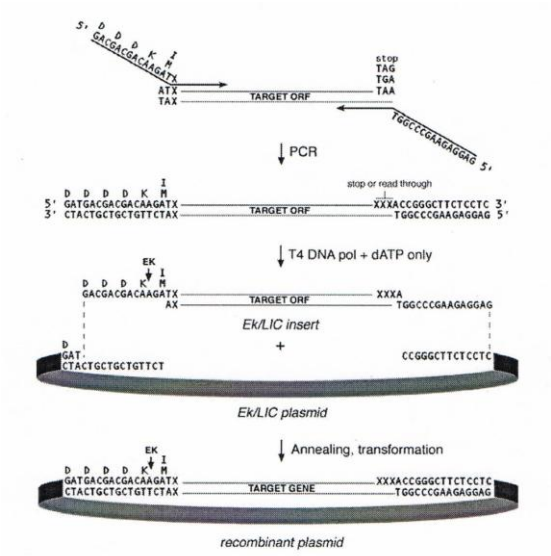


Figure 2.3 Strategy used for cloning TFP4 cDNA into the pET46 Ek/LIC vector. A PCR reaction with primers specifically designed to take advantage of the 5' and 3' overhangs in the vector is performed in the first step. T4 DNA Polymerase treatment of the PCR product creates compatible overhangs in the insert DNA. Final step is the annealing reaction resulting in a recombinant plasmid containing the DNA of interest (from BD Biosciences).

The T4 DNA polymerase has specific 3'-5' exonuclease activity which creates a very specific 13 to 14 single stranded overhangs in the vector and the selected DNA to be inserted. The plasmid contains an ampicillin resistant gene for growth in selected media, a 6xHisTag for purification using Ni charged columns and a *lac*UV5 promoter for T7 polymerase activation. Protein production is induced by isopropyl-1-thio-β-D-galactopyranoside (IPTG) which acts on the *lac*UV5 promoter. The annealed vector/insert was transformed into NovaBlue cells for characterization. A final transformation was then performed into *E.coli* strain BL21(DE3) cells for protein production and purification.

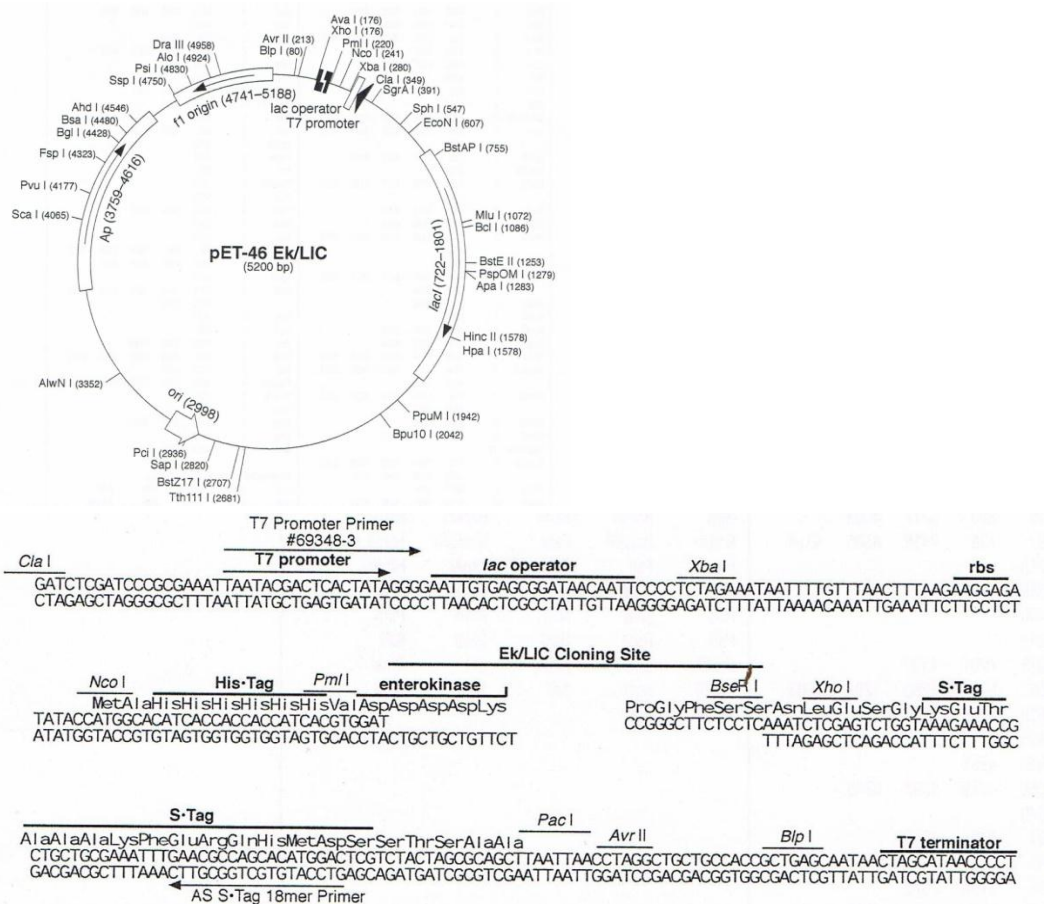


Figure 2.4 Map of pET46 Ek/LIC and sequence of cloning site. The vector incorporates a T7 promoter site and an adjacent downstream *lac* operator site, an enterokinase cutting site between the His-Tag and the protein sequence for purification of the protein without the tag; an Ampicillin resistance gene for growth in selective media (from BD Biosciences).

The resultant plasmid was transformed into NovaBlue competent cells by heat shock. This cell strain is suited as an initial cloning host because of its high transformation efficiency. The most important features are the *recA* and *endA* mutations. *RecA* is an *E.coli* protein essential for the repair and maintenance of DNA. By mutating this allele, the stability of extrachromosomal DNA is assured so the transformants will retain the DNA throughout cell divisions (Little *et al.,* 1984). *EndA* is a periplasmic nuclease and

inactivating it assures the efficiency of DNA uptake during transformation (Puyet *et al.,* 1990).

To express the TFP4 protein under the *lac*UV5 promoter, E.coli strain BL21(DE3) cells were transformed with the pET46TFP4 vector. These cells are engineered to take advantage of several features. The T7 RNA polymerase is under the control of the *lac*UV5 promoter which is present on the pET46 construct. It carries the *lacI* gene and LacI represses the expression of the T7 polymerase (Studier and al., 1990).The polymerase activity can be induced by isopropyl-1-thio-b-D-galactopyranoside (IPTG) which acts on the *lac*UV5 promoter and starts transcription. The cells also have mutated *lon 8* and *ompT* genes. The *lon 8* gene encodes a protein which regulates gene expression by targeting specific regulatory proteins for degradation (Wang et al., 1994). The OmpT protease is a membrane bound protein in wild type E. coli, presumably helping the cells to derive amino acids (Nakata *et al.,* 1993).

Materials and Methods

Isolation of Total mRNA from *C. formosanus* Shiraki Termite Soldiers and cDNA Library Construction

Between 150-200 healthy *C. formosanus* Shiraki termite soldiers were collected from colonies obtained from field sites in New Orleans, Louisiana. The frontal gland secretion was isolated on ice by squeezing the head near the frontal pore with tweezers and the secretion were transferred into 5 ml standard PBS buffer pH 7.0 and micro centrifuged at 7000 rpm for 10 min. The supernatant fraction was transferred to a new tube and filtered with a 2µm filter prior to dialysis for 48 hrs at a cut-off molecular weight of 3000 Da followed by lyophilization. The lyophilized powder was made up to a solution of 40 mg/ml with PBS buffer and stored at -20^{0}C.

For isolation of total mRNA from termite soldiers of *C. formosanus* we used the Fast Track 2.0 kit (BD Biosciences Clontech, Carlsbad Ca). A total amount of 1g soldier termite tissue was frozen in liquid nitrogen and lysed in 15 ml of lysis buffer (200 mM NaCl, 200 mM Tris-HCl, 1.5 mM $MgCl_2$, 2% Sodium Dodecyl Sulfate (SDS), 0.4 mg/ml Proteinase K, pH 7.5). The cell lysate was incubated for 60 minutes at 45^0C, centrifuged 5 minutes at 4000xg and the supernatant fraction collected. The supernatant fraction was mixed with 75 mg lyophilized oligo(dT) cellulose for 60 minutes and the mixture was pelleted 5 minutes at 3000xg. The supernatant fraction was removed and the resulting resin was resuspended in 20 ml binding buffer (500 mM NaCl, 10 mM Tris-HCl, pH 7.5), centrifuged 5 minutes at 3000xg and resuspended again in 10 ml of low salt buffer (250 mM NaCl, 10 mM Tris-HCl, pH 7.5) in order to remove contaminating RNAs, such as ribosomal RNAs. The total mRNA bound to the resin was eluted with 10 mM Tris-HCl, pH 7.5 and precipitated with 60 µl of 2M sodium acetate and 1 ml 200 proof ethanol. The resulting pellet was air dried and resuspended in 50 ul of10 mM Tris-HCl, pH 7.5. The final mRNA concentration was determined using a spectrophotometer by measuring light absorption at 260 nm and using the formula [RNA] = (A_{260}) x (0.04µg/µl) x D where D is the dilution factor.

Oligonucleotides used for cDNA library construction: 3'-primer mix: 5'-AAGCAGTGGTATCAACGCAGAGTAC$(T)_{30}$VN-3', V = A, G or C, first strand buffer: 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 6 mM $MgCl_2$, BD PowerScript^{TM} Reverse Transcriptase, 20mM Ditiothreitol(DTT), dNTP Mix (10mM each of dATP, dCTP, dGTP, dTTP), Tricine EDTA buffer (10mM Tricine-KOH, 1 mM EDTA).

RT reaction mix: 3 µl total mRNA, 1 µl 3'-primer, 1 µl dH_2O. This mix was incubated for 2 minutes at 70^0C then cooled on ice for 2 minutes. After incubation the following reagents were added: 2 µl 5x First Strand Buffer, 1 µl DTT, 1 µl dNTP mix, 1 µl BD PowerScript Reverse Transcriptase. The reaction mix was incubated for 1.5 hours at 42^0C in a water bath. Subsequently 20 µl Tricine EDTA buffer was added and the reaction mix was heated for 7 minutes at 72^0C. The sample was stored at -20^0C. The final product was a 3' RACE ready cDNA library.

TFP4 cDNA Isolation and Amplification

5'-primer mix: 5'-GAYTGBCARCTNTTYTGBCCNATG-3' where Y = T or C, B=T, C or G, R = A or G, 3'-primer mix: 5'-CTAATACGACTCACTATAGGGCAAGCAGTGG TATCAACGCAGAGT-3', where V=A, G or C and 5'-CTAATACGACTCACTATAGGGC-3', BD Advantage 2 PCR Buffer, dNTP Mix (10mM each of dATP, dCTP, dGTP, dTTP), BD Advantage 2 Polymerase mix.

PCR Reaction mix: 2.5 µl 3'RACE ready cDNA, 5 µl 3' primer mix, 1 µl 5' primer mix, 6 µl 10x BD Advantage 2 reaction buffer, 1 µl dNTP mix, 43.5 µl nuclease free water and 1 µl BD Advantage 2 Polymerase mix. For positive control 1 µl of the provided RACE TFR control primer was substituted for the 3' primer and the negative controls contained only one primer, either the 5' or the 3' primer.

PCR conditions: The reactions were performed using a Perkins Elmer DNA Thermal Cycler. The steps were as follows: 30 seconds at 94^0C, 3 minutes at 72^0C for 5 cycles; 30 seconds at 94^0C, 30 seconds at 70^0C, 3 minutes at 72^0C for 5 cycles; 30 seconds at 94^0C, 30 seconds at 68^0C, 3 minutes at 72^0C for 25 cycles; 10 minutes at 72^0C for 1 cycle.

PCR purification was performed using the QIAquick PCR Purification kit using the reagents provided and the DNA was eluted in 40 µl nuclease free water.

Plasmid Subcloning and Transformation into *E.coli* Strain BL21(DE3) Cells

Oligonucleotides used: 5' primer: 5'-GACGACGACAAGATAGAAGACTGCCAGC TGTT-3', 3' primer: 5'-GCTACACCAGCAACATCAACCGGGATTCTCCTC-3', BD Advantage 2 PCR Buffer, dNTP Mix (10 mM each of dATP, dCTP, dGTP, dTTP), BD Advantage 2 Polymerase mix.

PCR reaction mix: 2 µl DNA, 10 µl each 5' and 3' primers, 4 µl $MgCl_2$,10 µl 10x BD Advantage 2 reaction buffer, 10 µl dNTP mix, 53.5 µl nuclease free water and 0.5 µl BD Advantage 2 Polymerase mix.

PCR conditions: The reactions were performed using a Perkins Elmer DNA Thermal Cycler. The steps were as follows: 4 minutes at 94^0C for 1 cycle; 1 minute at 94^0C, 1 minute at 55^0C; 1 minute at 72^0C for 30 cycles; 10 minutes at 72^0C for 1 cycle.

PCR purification was performed using the QIAquick PCR Purification kit using the reagents provided and the DNA was resuspended in TlowE buffer (10 mM Tris-HCl and 0.1 mM EDTA pH 8.0).

T4 DNA Polymerase treatment was performed to generate 13 to 14 nucleotides overhangs in the DNA insert. The reaction mix contained 1 µl DNA insert, 2 µl 10x T4 DNA Polymerase buffer, 2 µl 25 mM dATP, 1 µl 100 mM DTT, 0.4 µl T4 DNA Polymerase (2.5U/ µl), 12.6 µl nuclease free water for a total volume of 20 µl. The mix was incubated 30 minutes at 22^0C and the enzyme was inactivated by heating at 75^0C for 20 minutes. The annealing reaction contained 1 µl pET 46Ek/LIC vector, 2 µl treated DNA insert and 1 µl 25 mM EDTA. The reaction was incubated at 22^0C for 60 minutes.

50 ul NovaBlue competent cells, Super Optimal broth with Catabolite repression (SOC) medium: 20 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 10mM $MgCl_2$, 10 mM $MgSO_4$, 0.4 g/l glucose; LB agar plates with ampicillin: 10g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract, 20 g/L agar; 50 mg/L ampicillin pH 7.0

One microliter DNA solution was added to 50 µl of cell suspension and the mix placed on ice for 5 minutes. The tubes were placed in a water bath at 42^0C for 30 seconds and immediately placed on ice for an additional 2 minutes. The solution was added to 250 µl of SOC medium and incubated at 37^0C with shaking at 250 rpm for 60 minutes. Fifty microliters of cells were plated on LB agar plates containing 50 mg/L ampicillin. Six colonies were picked for screening. A PCR reaction was performed as described before to check for the presence of the DNA insert.

100 µl *E.coli* BL21(DE3) cells, 1.42 M β-Mercaptoethanol (BME), SOC medium: 20 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 10mM $MgCl_2$, 10 mM $MgSO_4$, 0.4 g/l glucose, preheated at 42C; Luria Broth (LB) agar plates with ampicillin: 10g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract, 20 g/L agar; 50 mg/L ampicillin pH 7.0.

Before the transformation, 1.7 µl BME was added to 100 µl BL21(DE3) cells and the mix was incubated on ice for 10 minutes. Two microliters DNA solution containing the pET46TFP4 construct were added to the BL21(DE3) cells and the tubes placed on ice for 30 minutes. The tubes were then placed in a water bath at 42^0C for 45 seconds and immediately placed on ice for an additional 2 minutes. The solution was added to 900 µl of SOC medium and incubated at 37^0C with shaking at 250 rpm for 60 minutes. Fifty microliters of cells were plated on LB agar plates containing 50 mg/L ampicillin. Four colonies were picked for screening. A PCR reaction was performed as described before

in order to check for the presence of the DNA insert. The PCR product was then sequenced at GeneLab, on LSU's campus.

Results and Conclusions

As a general strategy, total mRNA was extracted from 1 g of *Coptotermes formosanus* Shiraki soldier tissue using the Fast track 2.0 kit from Clontech, a cDNA library was created, the TFP4 cDNA isolated using the BD SMART RACE kit, the insert cloned into pET46Ek/LIC vector and the plasmid transformed into *E.coli* BL21(DE3) cells.

Following mRNA extraction a reverse transcription reaction produced hybrid RNA/DNA which was used in PCR reactions containing specific and non-specific primers. The result was the isolation of TFP4 cDNA, with an expected size of 228 bp (figure 2.5).

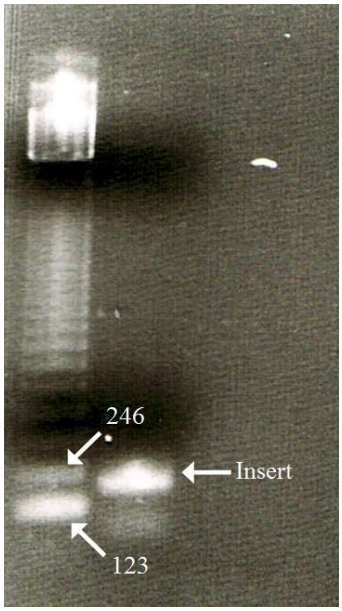


Figure 2.5 TFP4 DNA purified from PCR reaction using cDNA library.

The next step involved the cloning of TFP4 DNA using the pET46Ek/LIC system. A T4 DNA polymerase treatment was used to create 13-14 nucleotide overhangs in the DNA insert which were compatible with vector overhangs. An annealing reaction with the

vector followed, resulting in TFP4 DNA incorporation in the recombinant plasmid. This plasmid was transformed in NovaBlue cells by heat shock. The plates were overlaid with IPTG/X-Gal for blue/white screening (figure 2.6).

Blue colonies had an active form of β-galactosidase and were able to cleave x-galactose forming a blue pigment, therefore showing no insert was present in the plasmid. In the white colonies x-galactose was not hydrolyzed which indicates the presence of an insert because the *lacZ* gene was interrupted.

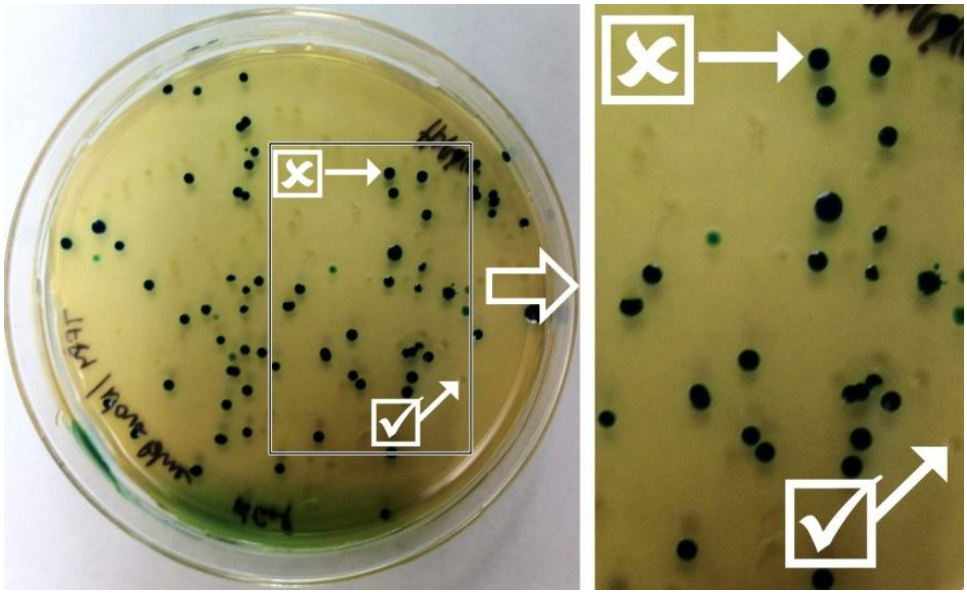


Figure 2.6 NovaBlue cell plate blue/white screening. White cells were successfully transformed while blue cells had the intact *lacZ* gene.

Plasmids were isolated from NovaBlue cells and transformed into BL21(DE3) cells by heat shock (figure 2.7). In order to verify the presence of a successfully cloned plasmid, BL21(DE3) cells were grown overnight and the plasmid isolated.

DNA was cut with NcoI and XhoI restriction endonucleases (figures 2.8a and 2.8b). In the original plasmid the expected size of the cut DNA fragment was 65 bp and in the cloned plasmid the expected size was 293 bp. The pET46Ek/LIC size is 5200 bp.

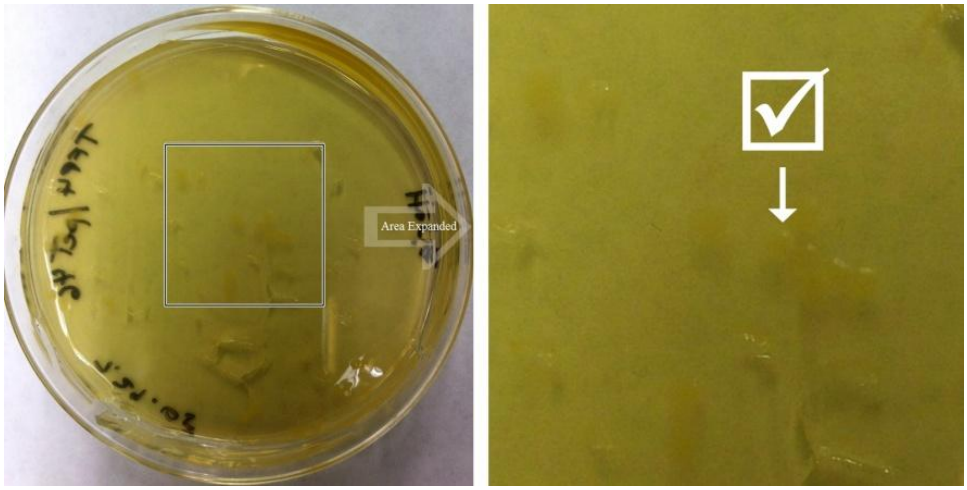


Figure 2.7 BL21(DE3) cells after transformation with pET46TFP4 plasmid.

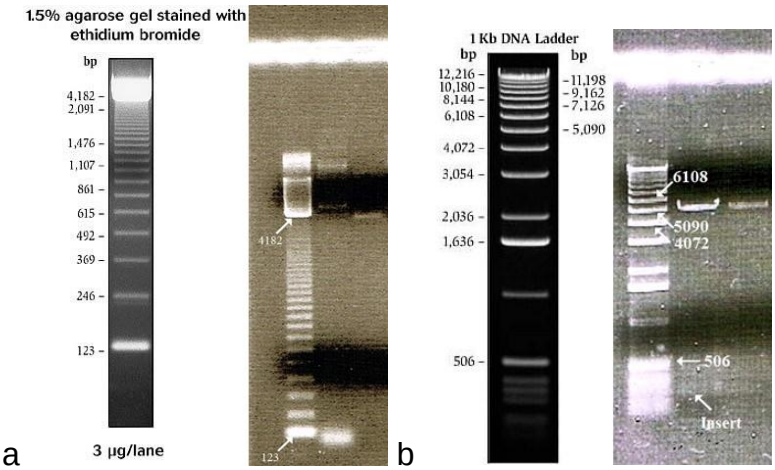


Figure 2.8 a. pET46 Ek/LIC with no insert, cut with NcoI and XhoI restriction endonucleases; b. pET46TFP4 cut with NcoI and XhoI restriction endonucleases.

The recombinant plasmid was sent for sequencing to GeneLab, on LSU campus, and the results are presented in figure 2.9.

TFP4 DNA sequence – 228 bases

ATGGCACATCACCACCACCATCACGTGGATGACGACGACAAGCCAGAAGACTGCCAGCTGTTCTGCCCAA TGATCTATGCTCCAATATGCGCCACAGACGGCGTCAGCCAAAGAACCTTCAGCAATCCATGTGACCTGAA AGTTTATAACTGCTGGAACCCGGATAATCCCTACAAAGAAGTGAAAGTAGGTGAATGTGATGATGCGAATA AACCTGTACCAATCTGA

TFP4 Protein sequence – 75 AA

MAHHHHHHVDDDDKPEDCQLFCPMIYAPICATDGVSQRTFSNPCDLKVYNCWNPDNPYKEVKVGECDDANKP VPI.

Red – sequence from vector pET46 Ek/LIC

Black – inserted DNA

Figure 2.9 DNA and amino acid sequences of TFP4

CHAPTER THREE: TFP4 PROTEIN EXPRESSION AND PURIFICATION

Introduction

A polyhistidine tag is an amino acid motif in proteins which consists of at least five histidine residues. Polyhistidine tags are used for affinity purification of recombinant proteins expressed in various prokaryotic systems. As a general strategy, bacterial cells grown under specific conditions are harvested by centrifugation and lysed, generally by mechanical means or by the use of detergents. The lysate, which is a mix of recombinant and host proteins, is incubated with affinity media containing bound nickel or cobalt ions to which the polyhistidine tag binds with micromolar affinity. The column is washed with a buffer to remove the unbound proteins resulting in relatively pure recombinant proteins (Hengen *et al.,* 1995). The His-tag is usually removed with enterokinase. In TFP4, initially we considered the protein a lysozyme because of its activity. We had problems removing the tag with enterokinase and it led us to test the protein for protease inhibitor activity.

Materials and Methods

TFP4 Protein Expression

BL21(DE3) cells carrying the pET46TFP4 construct were grown at 37^0C in 100 ml LB medium (10g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract, 50 mg/L ampicillin pH 7.0). In the exponential growth phase, the culture was diluted 1:10 with LB medium and incubated for another hour at 37^0C. Protein expression was induced with 0.3mM IPTG. After further cultivation for 8 hours at 37C the cells were harvested by centrifugation for 10 minutes at 10,000xg using a Sorvall^R RB-5G centrifuge. The supernatant fraction

was discarded and the cell pellet resuspended in BugBuster™ protein extraction reagent (Novagen) at a ratio of 1g cells per 5 ml reagent. Additionally, 25U Benzonase nuclease (Novagen) per ml was added to disrupt the cell DNA. The lysate was mixed for 25 minutes at room temperature and cellular debris was removed by centrifugation for 20 minutes at 14,000xg. The supernatant fraction was stored at 4^0C if the purification was done the same day or at -20^0C overnight.

TFP4 Protein Purification

Column charging buffer: 50 mM $NiSO_4$; column binding buffer: 0.5 M NaCl, 20mM Tris-HCl, 5 mM imidazole pH 7.9; column wash buffer: 0.5 M NaCl, 20mM Tris-HCl, 60 mM imidazole pH 7.9; elution buffer: 0.5 M NaCl, 20mM Tris-HCl, 1 M imidazole pH 7.9.

The His-Bind® resin was added to the column and the following sequence of washes was used to charge and equilibrate the column: 3 bed volumes sterile dH_2O, 5 volumes charging buffer, 3 volumes binding buffer. After the last wash was drained, the prepared extract was loaded to the top of the column. A flow rate of 10 bed volumes per hour was used. The column was then washed with 10 bed volumes of binding buffer followed by 6 volumes of wash buffer. The recombinant protein was eluted in 750ul aliquots using the elution buffer.

Protein concentrations were estimated from absorbance at 280 nm using a Beckman DU 530 spectrophotometer. The molar absorption coefficient was calculated to be 9970 cm^{-1} M^{-1} based on the formula $e_{280}(cm^{-1} M^{-1}) = (\#Trp) \times 5500 + (\#Tyr) \times 1490 + (\#Cys) \times 125$ (Pace *et al.,* 1995). The formula for protein concentration used was Concentration $_{[Protein]} = A_{280} / e_{280}$ (M) for a 1 cm cuvette.

SDS polyacrylamide gel electrophoresis (SDS-PAGE) of the purified protein in 12.5% polyacrylamide was conducted by the method of Laemmli (Laemmli *et al.,* 1970). A prestained marker (New England Biolabs) was used as a molecular weight marker. The apparent molecular weight is indicated: 175 kDa Maltose-binding protein (MBP)-b-galactosidase, 83 kDa MBP-paramyosin, 62 kDa Glutamic dehydrogenase, 47.5 kDa Aldolase, 32.5 kDa Triosephosphate isomerase, 25 kDa b-lactoglobulin A, 16.5 kDa Lysozyme, 6.5 kDa Aprotinin.

Results and Conclusions

The purified TFP4 fractions were loaded into a 12.5% polyacrylamide gel (figure 3.1) and the electrophoresis was run at 175 V until the dye front reached the bottom of the gel. The first lane contained the prestained protein marker from New England Biolabs, lanes 2 through 7 were loaded with 10 µl sample taken from each of the first six fractions eluted from the purifying Ni column, lane 8 had 10 µl cell lysate taken after the addition of BugBuster protein extraction reagent (first step of processing), lane 9 had 10 µl of cell lysate taken right before loading in the column (last step of processing). Lane 10 contained 10 µl of crude protein extract from termite soldiers reconstituted in PBS pH 7.0 at 40 mg/ml.

The theoretical molecular weight of TFP4 protein was calculated at 8.5 KDa and is expected to migrate slightly above the 6.5 KDa band of the standard, which is what can be inferred from the gel in figure 3.1. Also, judging from the band intensities, the Ni column method appears to achieve over 90% purity of the His-tagged TFP4 protein.

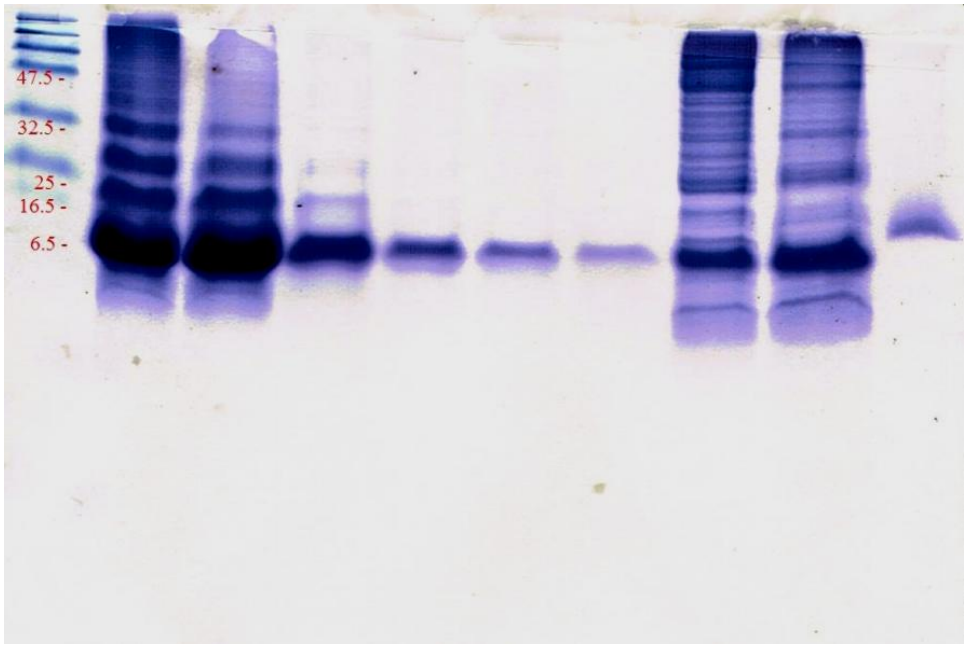


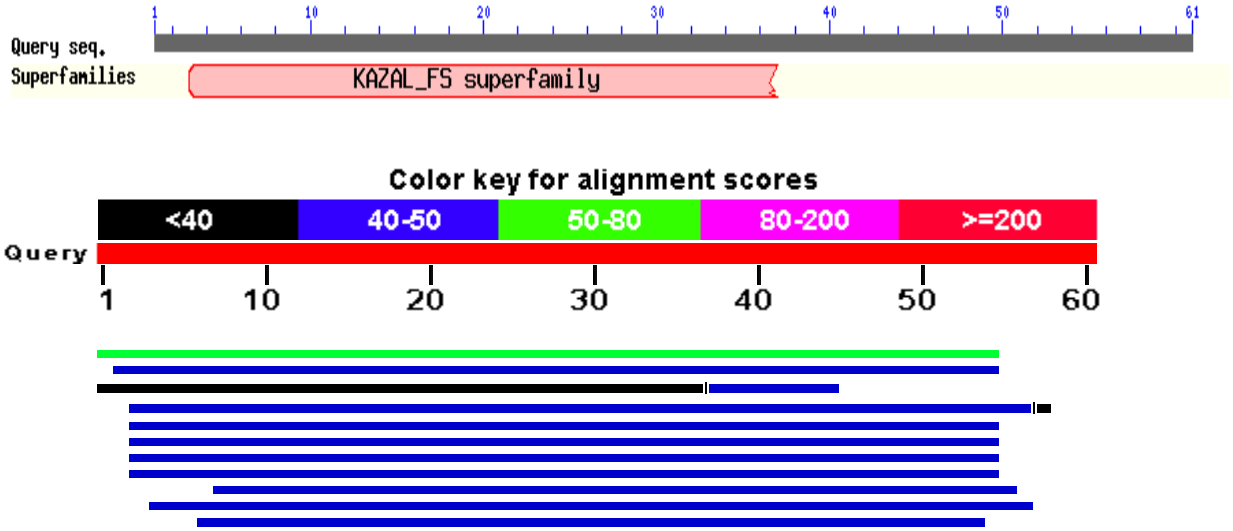
Figure 3.1 Fractions of purified TFP4 run on 12.5% SDS-PAGE. Lane 1 - protein standard, lanes 2-7 - eluted fractions from Ni column, lanes 8-9 - cell lysate at various stages of processing, lane 10 - crude protein extract from gland secretion. All lanes contain 10 µl of sample and 10 µl of loading buffer.

The aim of this process was to identify, clone and express the protein in the defense gland secretion of *C.formosanus* Shiraki termite soldiers corresponding to the N-terminal sequence EDC/WQLFC/WPMIY.

In conclusion, total mRNA was obtained from termite soldier tissue and the DNA of interest identified, amplified and cloned in an expression vector (pET46Ek/LIC). Following transformation into NovaBlue cells, the recombinant plasmid was cut with specific endonucleases to verify the insert and transformed into BL21(DE3) cells for the expression of the cloned protein . The DNA sequence was obtained and the expressed protein was purified on affinity Ni columns for further investigations of its activity.

A BLASTp search was conducted using the protein sequence obtained and the results summarized in figure 3.2.

Putative conserved domains have been detected, click on the image below for detailed results.



Accession	Description	Max score	Total score	Query coverage	E value
ADB02828.1	salivary glands proteinase inhibitor [Nauphoeta cinerea]	53.5	53.5	90%	6e-09
ADB02827.1	salivary glands proteinase inhibitor [Nauphoeta cinerea]	49.7	49.7	88%	3e-07
XP_001954335.1	GF16798 [Drosophila ananassae] >gb|EDV42896.1| GF16798 [Drosophila ananassae]	45.1	77.4	73%	5e-05
AAQ22771.1	serine proteinase inhibitor [Procambarus clarkii]	45.8	193	91%	5e-05

Figure 3.2 Sequences producing significant alignments and sequence alignments between TFP4 and salivary glands proteinase inhibitor from *Nauphoeta cinerea*; TFP4 and serine proteinase inhibitor from *Procambarus clarkii.*

A Kazal domain was identified using BLASTp software which analyzes similarities with known sequences deposited in NCBI database (figure 3.3). The Kazal domain is the inhibitory domain of serine proteases, a group of enzymes which includes trypsin, chymotrypsin and elastase.

```
gb|ADB02828.1|  salivary glands proteinase inhibitor [Nauphoeta cinerea]
Length=72

Identities = 22/55 (40%), Positives = 30/55 (55%), Gaps = 0/55 (0%)

           MEDCQLFCPMIYAPICATDGVSQRTFSNPCDLKVYNCWNPDNPYKEVKVGECDDA  55
            M +C + C    Y P+C  D     TF N C    YNC +PD   + V++GEC+DA
           MAECDIACTREYDPVCGADATHAETFGNACMFVFYNCQHPDAMMRLVRLGECNDA  72


gb|AAQ22771.1|  serine proteinase inhibitor [Procambarus clarkii]
Length=277

Identities = 23/55 (42%), Positives = 29/55 (53%), Gaps = 3/55 (5%)

            DCQLFCPMIYAPICATDGVSQRTFSNPCDLKVYNCWNPDNPYKEVKVGECDDANK  57
              C   CP IY P+C TDG   +T+SN C L V  C NP    +    GEC  +N+
            QCNSVCPQIYQPVCGTDG---KTYSNQCTLDVAACNNPQLHLRTAYQGECRTSNQ  225
```

Figure 3.3 Sequences producing significant alignments

The first sequence identified was from *Nauphoeta cinerea,* also known as Lobster Roach, a cockroach species found in tropical climates. This is not surprising since termites and cockroaches descend from a common ancestor. The results showed 40% identity (22 of 55 identical amino acids), 55% functional identity between the two proteins (30 of 55 amino acids identical or interchangeable) and a length of 72 amino acids which is comparable to TFP4.

The serine protease inhibitor from *Procambarus clarkii* (freshwater crawfish) had a 42% identity and 53% functional identity, but it also had a 3 amino acid gap and it was much larger, 277 residues long.

CHAPTER FOUR: DETERMINATION OF TFP4 LYSOZYME ACTIVITY

Introduction

“In this communication I wish to draw attention to a substance in the tissues and secretions of the body, which is capable of rapidly dissolving certain bacteria. As this substance has properties akin to those of ferments I have called it a “Lysozyme” and shall refer to it by this name throughout the communication” (Fleming, 1922), wrote Alexander Fleming in a paper where the discovery of lysozyme was announced and the term was coined. Lysozymes are low molecular weight enzymes which hydrolyze β (1,4) glucosidic linkages between N-acetyl-muramic acid (NAM) and N-acetyl-D-glucosamine (NAG) residues present in the mucopolysaccharide cell wall of a variety of microorganisms (Helal *et al.,* 2010).

Lysozyme is part of the innate immune system and is present in various mucosal secretions, such as tears, saliva and mucus. Its antibiotic activity is one of the first lines of defense against bacteria in metazoans. More than 80 years of research on lysozymes have shown several genetic families of lysozymes with impressive variety, and the enzyme has become a model for enzymology (Jolles, 1997). Lysozyme was one of the first enzymes to be completely sequenced (Canfield, 1963) and the first enzyme structure to be solved by X-ray crystallography with 2Å resolution (Blake, 1965).

The Phillips catalytic mechanism was widely accepted for the model lysozyme from chicken egg white; however, recent evidence led to the revision of the Kornberg mechanism (Koshland *et al.,* 1953). According to Phillips mechanism the lysozyme binds to a hexasaccharide and distorts the fourth sugar into a half-chair configuration.

Glutamate acts as a proton donor to the glycosidic oxygen and aspartate acts as a nucleophile, stabilizing the glycosyl-enzyme complex through electrostatic interactions. The C-O bond is cleaved and interaction with a water molecule yields the hydrolyzed sugar.

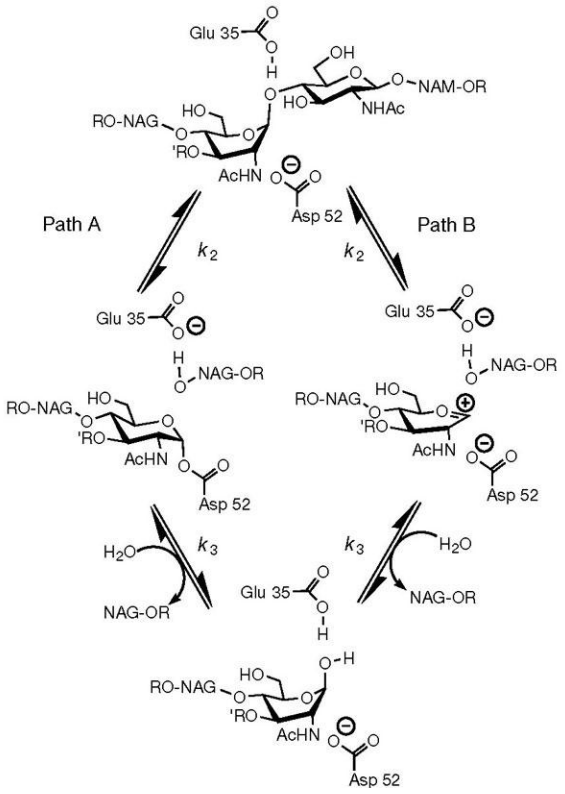


Figure 4.1 Catalytic mechanism of a lysozyme (from Vocadlo *et al.,* 2001)

Recently, though, based on experimental data from mass spectrometry, crystal structures, kinetic isotope effects and enzyme kinetics the Kornberg mechanism was revised with evidence of a covalent glycosyl-enzyme intermediate complex instead of the oxocarbenium ion in the Phillips mechanism (Vocadlo *et al.,* 2001). Our group contributed to this study (figure 4.1).

Zymograms have been used extensively in the determination of enzymatic activity of lysozymes and other cell wall hydrolyses (Potvin *et al.,* 1988; Audy *et al.,* 1989; Leclerc *et al.,* 1989). The method is based on the separation of proteins by gel electrophoresis

followed by a refolding step where the proteins are renatured and the enzymes are capable of converting the substrates that have been incorporated into the gel, overlaid or spread onto the gel. The enzymes are separated by electrophoresis under denaturing conditions which prevent the interaction with the substrate. In the renaturing step, the enzymes are allowed to refold and hydrolyze the embedded substrate resulting in clear zones in the opaque gels. This method allows for the simultaneous detection of enzyme activity and estimation of molecular weight. The assay, developed in our laboratory, uses *Micrococcus lysodeikticus* cells embedded in the gel and stained with the vinyl-sulfone reactive dye Remazol brilliant blue R (RBB) reagent (Hardt *et al.,* 2003). RBB was shown to bind the hydroxyl group of sugars (Stamm, 1963) which are present in peptidoglycan, the major constituent of bacterial cell wall. The hydrolysis of *M. lysodeikticus* cell wall stained with RBB leads to the release of soluble blue product and the presence of a clear zone where the activity occurred.

Another method of determining enzymatic activity is to measure it in solution. *M. lysodeikticus* cell walls may be labeled with a fluorescent dye. Enzymatic activity releases the dye in solution where fluorescence can be measured using a spectrofluorometer.

The assay relies on lysozyme activity on *M. lysodeikticus* cell walls labeled with a fluorescent dye such that fluorescence is quenched. Hydrolysis of glucosidic linkages relieves the quenching, resulting in an increase in fluorescence proportional to enzymatic activity. The fluorescence increase can be measured using a spectrofluorometer that can detect fluorescein. Digestion products from the substrate have absorption maxima at 494 nm and fluorescence emission maxima at 518 nm.

Materials and Methods

Lysozyme Enzymatic Activity Determined by Zymogram

Materials required for gel preparation, lyophilized *M. lysodeikticus* cells, chicken egg white lysozyme, bovine serum albumin, RBB were purchased from Sigma-Aldrich Co.

M. lysodeikticus whole cells were labeled with RBB according to the procedure for the synthesis of RBB-starch (Rinderknecht *et al.,* 1967) with modifications suggested by Ito (Ito *et al.,* 1992). A solution of 200 mg of RBB-R in 80 ml distilled water was added to a suspension of 300 mg *M. lysodeikticus cells* in 80 ml distilled water at 50^0C. The mixture was stirred 30 minutes when 4g of Na_2SO_4 and 200 mg Na_3PO_4 was added. After centrifugation for 10 minutes at 2600 x g the cell pellet was resuspended and the supernatant discarded. This step was repeated until the supernatant became clear. The cells were washed with distilled water and dried for lyophilization.

SDS-PAGE was performed in 12.5% polyacrylamide gel which contained 0.1% (w/v) labeled *M. lysodeikticus* cells. The gel and buffers contained 0.1% SDS. The samples of 20 µl were boiled for 5 minutes in loading buffer (12% sucrose, 2% SDS, 50 mM Na_2CO_3, 0.1% bromophenol blue, 10% 1 M DTT) prior to loading in the gel. A prestained protein marker (New England Biolabs) was used as a molecular weight standard. Electrophoresis was performed using a MiniProtean II cell (Biorad) at a constant voltage of 175 V. The gel was washed twice with distilled water for 30 minutes to remove SDS and incubated with 300 ml refolding buffer (100 mM NaH_2PO_4, 100 mM Na_2HPO_4, 10% (v/v) Triton X-100 pH 7.0) .

Lysozyme Enzymatic Activity Determined by Fluorescence

M. lysodeikticus cells labeled with fluoroscein using reaction buffer (0.1 M Na_3PO_4, 0.1 M NaCl pH 7.5) from Molecular Probes Inc.

Hen egg white lysozyme (Sigma-Aldrich) was dissolved in refolding buffer at various concentrations and used as a positive control. The test protein was diluted in the provided reaction buffer at various concentrations. The substrate was used at a concentration of 50 µg/ml. Fifty microliters of protein solution and 50 µl of substrate solution were mixed and the reaction was allowed to proceed one hour at 37C in the dark. The emission at 518 nm were detected using a SpectraMax fluoroscope and data was recorded and analyzed using SoftMax Pro software.

Results and Conclusions

Lysozyme Enzymatic Activity Determined by Zymogram

M. lysodeikticus cells were widely used as a substrate to measure lysozyme activity in spectrophotometric assays (Shugar, 1953) or radial diffusion assays (Osserman, 1966). The method used to assay the activity of *C. formosanus* protein TFP4 was developed by Markus Hardt (Hardt *et al.,* 2003). He assayed ostrich egg white lysozyme (figure 4.2a) and chicken egg white lysozyme (figure 4.2b) and interpreted the clear zone as being areas where enzymatic activity occurred. Chicken egg white lysozyme is a 14,500 Da protein while ostrich egg white lysozyme is a 20,500 KDa.

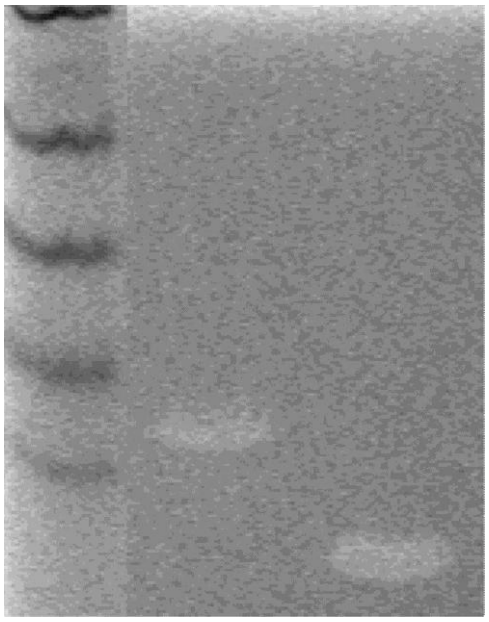
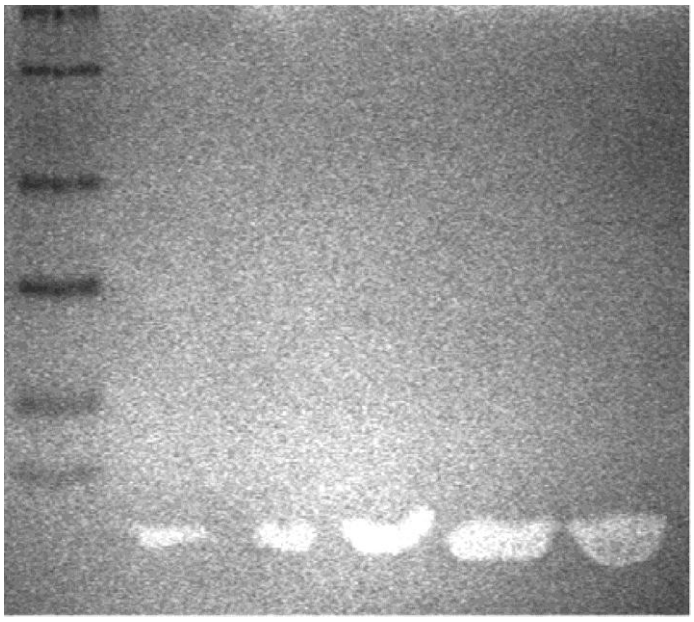
a.  b. 

Figure 4.2 Zymogram of ostrich egg white lysozyme and hen egg white lysozyme (a) and zymogram of hen egg white lysozyme at various concentrations - 25, 50, 125, 250 and 500 units (b) from Hardt et al., 2003

Figure 4.3 represents a zymogram with HEWL and TFP4 after incubation for 16 hrs in dark at 37^0C. In the lane where TFP4 was loaded there is a clear zone in the opaque gel, indicating lysozyme activity. The same clear zone is observed in the lane where HEWL was loaded, at a different molecular weight. In order to eliminate the possibility of sample contamination another zymogram was run where TFP4 was added alone or together with HEWL in the same sample (figure 4.4).

The same pattern of migration is observed, with TFP4 running lower than HEWL and it is clearly discernible that in the second lane there are two separate proteins which hydrolyze *M. lysodeikticus* cell wall.

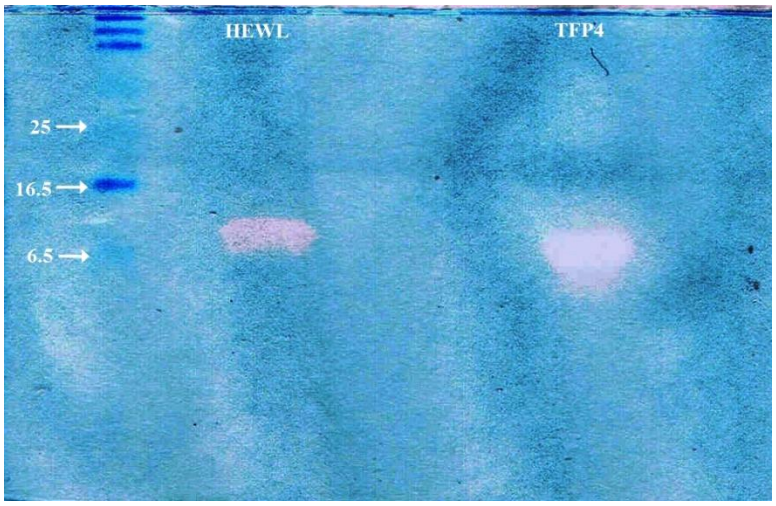


Figure 4.3 Zymogram of HEWL and TFP4 shown on 12.5% SDS-polyacrylamide gel embedded with 0.01% Remazol blue-labeled *M. lysodeikticus* cells.

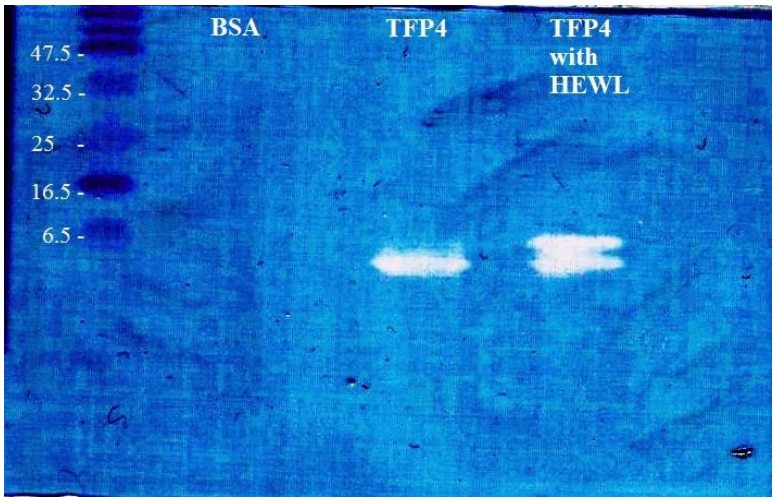


Figure 4.4 Zymogram of TFP4 and TFP4 mixed with HEWL. The proteins were denatured prior to loading and the gel was incubated in refolding sodium phosphate buffer for 16 hrs at 37C.

Lysozyme Enzymatic Activity Determined by Fluorescence

Chicken egg white lysozyme hydrolyses the glycosidic linkages between N-acetyl-muramic acid and N-acetyl-D-glucosamine residues in bacterial cell walls. In the assays performed *M. lysodeikticus* cells were labeled with fluorescein in such amount that fluorescence was quenched. The enzymatic activity of lysozyme hydrolyzed the

glycosidic bonds and residues labeled with fluorescein were released in solution where the fluorescent signal was detected. Figure 4.5 shows a fluorescent signal scan (in events per minute) between 500 nm and 650 nm. Three amounts of HEWL were used - 37.5 U, 75 U and 150 U. One unit (U) is defined as the amount of enzyme required to produce a change in the absorbance at 450 nm of 0.001 units per minute at 25^{0}C using a suspension of *M. lysodeikticus* cells as substrate. The graph shows enzyme activity which is dependent on the amount of enzyme in the reaction with levels between 20-28 million counts per minute.

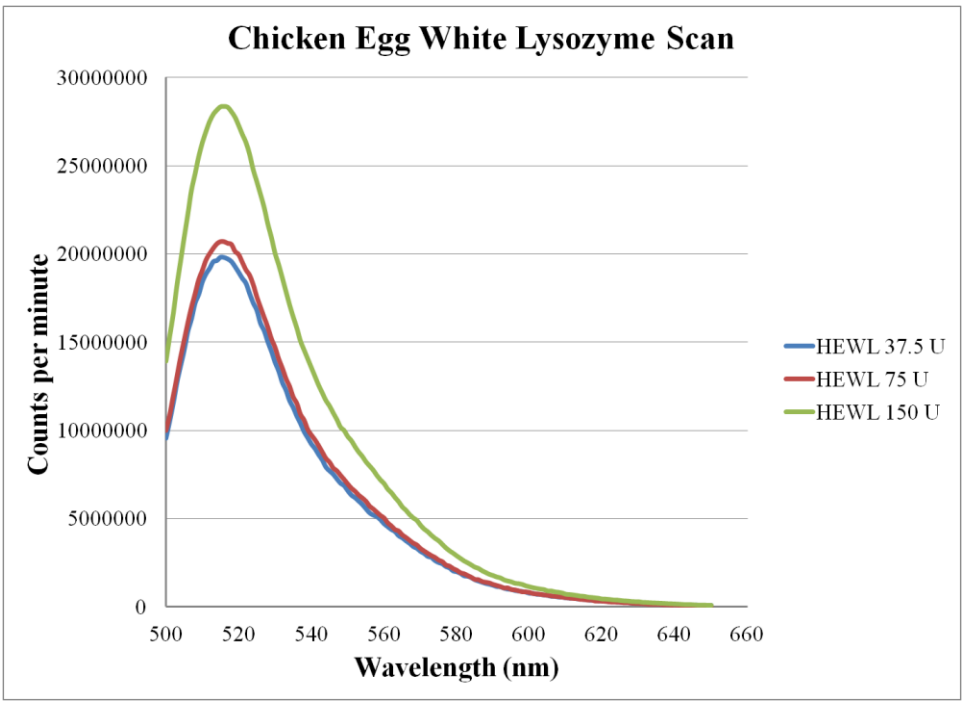


Figure 4.5 Fluorescence emission spectrum of chicken egg white lysozyme. Fifty microliters of HEWL solution and 50 µl of substrate solution at 50 µg/ml were mixed together and the reaction was allowed to proceed one hour at 37^{0}C in the dark.

Figure 4.6 shows the signal detected when the substrate was mixed with TFP4. Three amounts of protein were used - 12.5 µg, 25 µg and 50 µg. The graph shows lysozomal activity with maximum signals between 1.5-5.4 million counts per minute. The signal

recorded is lower than the HEWL signal but there are two possible explanations. One explanation is that TFP4's specific activity might be inherently lower than HEWL's; it might be a less efficient enzyme. A second plausible argument is the termite protein was expressed in a prokaryotic system and a proportion of it was not correctly folded so its enzymatic activity was mitigated by in an incorrect conformation. But the conclusion that can be drawn from the assays is that the termite protein functions as a lysozyme breaking the β-(1,4) linkages between N-acetyl-muramic acid and N-acetyl-D-glucosamine residues present in the mucopolysaccharide cell wall.

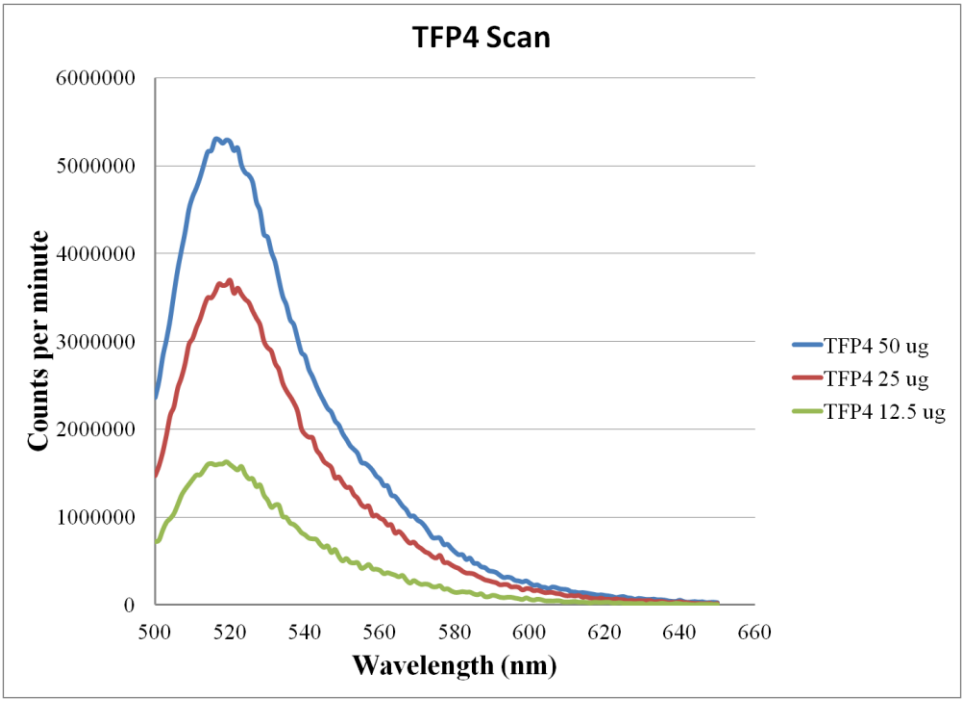
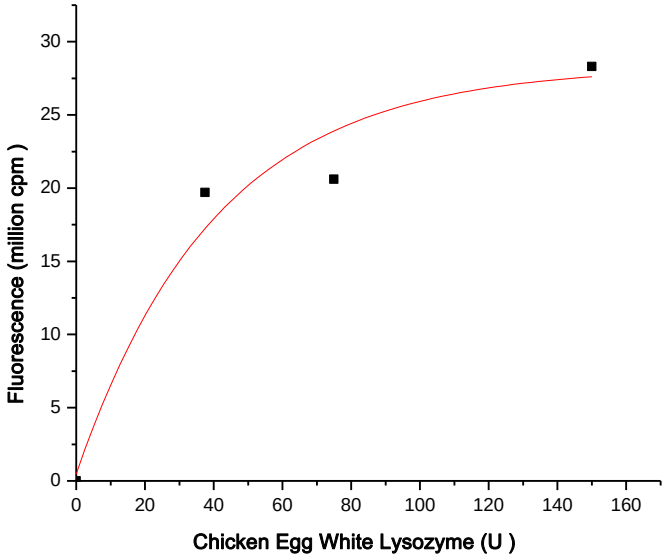
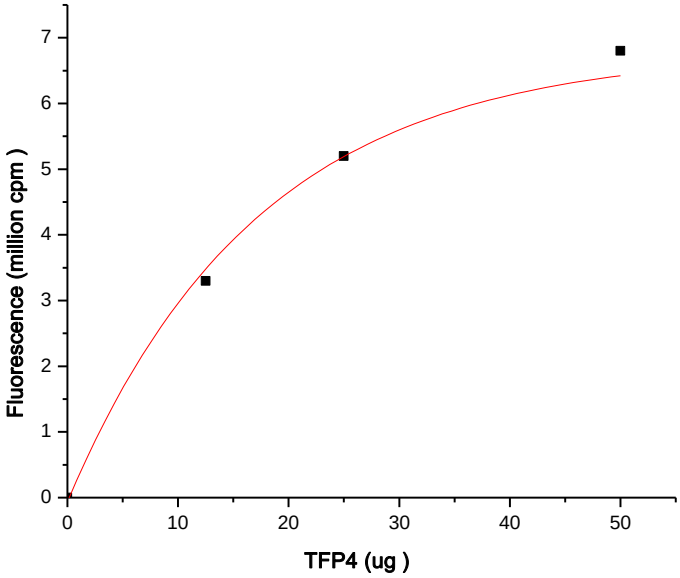


Figure 4.6 Fluorescence emission spectrum of TFP4. Fifty microliters of TFP4 solution and 50 µl of substrate solution at 50 µg /ml were mixed and the reaction was allowed to proceed one hour at 37C in the dark.

Figures 4.7a and 4.7b represent plots of the emission maxima at 518 nm for the HEWL and TFP4 assays. An exponential curve was fit to the data points with R^2 values of 0.96 for the HEWL assay and 0.99 for the TFP4 assay.



a.



b.

Figure 4.7 Plot of HEWL (a) and TFP4 (b) emission maxima at 518 nm.

The R^2 coefficient of determination is a statistical measure of how well the regression line approximates the real data points with values between 0 and 1. An R^2 of 1.0 indicates that the regression line perfectly fits the data. The shapes of both graphs indicate the enzymatic reactions have reached a plateau where no more substrate is available for hydrolysis.

Finally, the experimental data indicate the termite protein's profile is similar to HEWL's profile, meaning the TFP4 protein is a functional lysozyme. Fluorescence was observed in both assays as a function of the amount of protein in the solution, observation supported by the values of R^2, the statistical "goodness of fit" coefficient.

CHAPTER FIVE: DETERMINATION OF TFP4 SERINE PROTEASE INHIBITION ACTIVITY

Introduction

Serine proteases are found ubiquitously in eukaryotes and prokaryotes and are enzymes that cleave peptide bonds in proteins in which serine is the nucleophile amino acid at the active site.

MEROPS is a database of all known proteolytic enzymes based on statistically significant similarities in sequence and structure. This system categorizes proteins into clans based on catalytic mechanism and common ancestry (Rawlings et al., 2008). Serine proteases comprise over one third of all proteolytic enzymes and are grouped in 13 clans. Clan PA proteases are highly represented in eukaryotes and are the largest family of serine proteases (Cera, 2009).

Table 5.1. Serine proteases classification (from Cera, 2009)

Clan	Families	Representative member	Catalytic residues	Primary specificity
PA	12	Trypsin	His, Asp, Ser	A, E, F, G, K, Q, R, W, Y
SB	2	Subtilisin	Asp, His, Ser	F, W, Y
SC	2	Prolyl oligopeptidase	Ser, Asp, His	G, P
SE	6	D-A, D-A carboxypeptidase	Ser, Lys	D-A
SF	3	LexA peptidase	Ser, Lys/His	A
SH	2	Cytomegalovirus assemblin	His, Ser, His	A
SJ	1	Lon peptidase	Ser, Lys	K, L, M, R, S
SK	2	Clp peptidase	Ser, His, Asp	A
SP	3	Nucleoporin	His, Ser	F
SQ	1	Aminopeptidase DmpA	Ser	A, G, K, R
SR	1	Lactoferrin	Lys, Ser	K, R
SS	1	L,D-carboxypeptidase	Ser, Glu, His	K
ST	5	Rhomboid	His, Ser	D, E

The catalytic mechanism is preserved in all serine proteases and it consists of a catalytic triad located in the active site of the enzyme. The triad of amino acids – histidine, serine and aspartic acid - allows for the transfer of protons in and out of the enzyme's active site. (Carter *et al.,* 1988).

The subtilisin clan of proteases is found in prokaryotes and is an example of convergent evolution. It utilizes the same catalytic triad mechanism but structurally and evolutionary is unrelated to the eukaryotic group (Brenner *et al.,* 1988).

According to their protein structure, the serine proteases can be categorized in chymotrypsin-like and subtilisin-like proteases. The chymotrypsin group can be further separated by the substrate they hydrolyze in trypsin-like, chymotrypsin-like and elastase-like proteases. The trypsin-like enzymes cleave the peptide bond following a positively charged amino acid (lysine or arginine). Chymotrypsin-like enzymes have specificity for cleavage on the peptide bond following hydrophobic residues such as tyrosine, phenylalanine and tryptophan; and elastase-like proteases cleave the peptide bond following a smaller residue such as alanine glycine or valine (Rawlings *et al.,* 2008; Page *et al.,* 2008).

In the catalytic mechanism two tetrahedral intermediates are generated. In general, the enzyme binds the substrate at the catalytic site, the peptide bond is cleaved, the N-terminus region of the peptide is released, a water molecule binds the acyl-enzyme intermediate and the C-terminus and the water molecule are released. In this mechanism the serine –OH group acts as a nucleophile attacking the carbonyl carbon of the peptide bond, assisted by the nitrogen of histidine and the carbonyl group of the aspartic acid. As a result a tetrahedral intermediate is generated (figure 5.1). The peptide bond is broken and the result is an acyl-enzyme. Next a molecule of water attacks the carbonyl carbon coordinated by the nitrogen of histidine which accepts a proton from the water molecule generating a second tetrahedral intermediate. Finally the C-O bond formed in the first step between serine and the peptide is destabilized by

the proton accepted by the histidine residue and the C-terminus part of the initial peptide is ejected from the catalytic site along with a water molecule (Hedstrom, 2002).

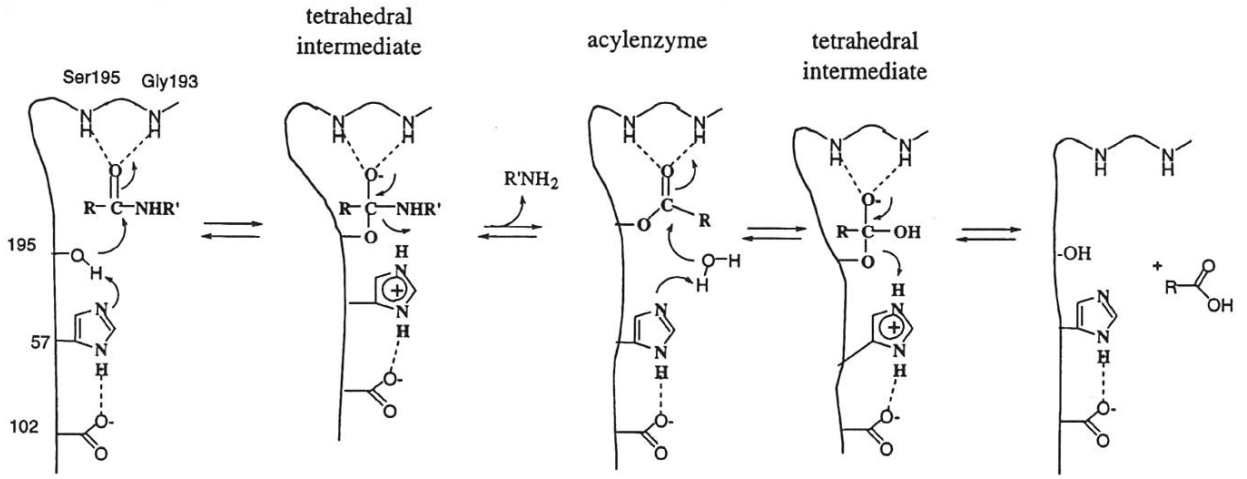


Figure 5.1 General mechanism of serine protease hydrolysis of peptide bond (from Hedstrom, 2002)

Proteases play a significant role in diverse physiological processes such as food digestion, blood clotting, embryogenesis, defense mechanisms and innate immunity. Many processes occur as proteolytic cascades leading to cellular responses and activation/inactivation of these events has to be controlled at different regulatory levels. Most animal species express a variety of protease inhibitors whose function in certain locations is to prevent unwanted proteolysis (Clynen *et al.,* 2005).

In invertebrates, serine proteases and their inhibitors are known to be involved in coagulation, complement activation (Iwaga 1993) and the innate immune response (Imler *et al.,* 2000). Pathogens also use protease inhibitors to suppress proteases that are an integral part of the host's defense (Tian *et al.,* 2005).

There is a wide variety in the length of amino acid sequences of the identified serine protease inhibitors, ranging from 29 to more than 400 residues. However, only two different mechanisms of inhibition have been proposed. For small proteins (29 to 190 amino acids), the cognate enzyme is bound in a substrate-like mechanism (Laskowski and Kato, 1980). The serpins (serine protease inhibitors) comprise a family of large glycoproteins which function as suicide substrate inhibitors (Bode *et al.,* 1992). Many serpins regulate enzymatic processes in human plasma, being involved in mitigation of inflammatory reactions, fibrinolysis and blood coagulation (Schoofs *et al.,* 2002).

Serine protease inhibitors can be classified in at least 10 homologous families (table 2). Members of the same family share a conserved topological relationship between the disulfide bridges and the reactive site (Laskowski and Kato, 1980).

Table 5.2. Trypsin inhibitors classification (from Laskowski and Kato, 1980)

I.	Bovine pancreatic trypsin inhibitor (Kunitz) family
II.	Pancreatic secretory trypsin inhibitor (Kazal) family
III.	*Streptomyces* subtilisin inhibitor family
IV.	Soybean trypsin inhibitor (Kunitz) family
V.	Soybean proteinase inhibitor (Bowman-Birk) family
VI.	Potato I inhibitor family
VII.	Potato II inhibitor family[a]
VIII.	*Ascaris* trypsin inhibitor family[a]
IX.	Other families

However, most inhibitors interact with their cognate enzyme according to a common mechanism. Each inhibitor has at least one reactive site (P_1) which specifically interacts with the enzyme's active site (table 3). The complex formation occurs with little conformational change – a lock and key mechanism. The P_1 reactive site is located at the protein's surface and is encompassed in at least one disulfide loop which ensures the hydrolyzed peptide chain cannot dissociate from the enzyme.

Table 5.3. Consensus sequences for the inhibitory site (from Laskowski and Kato, 1980)

	P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'	P_4'	P_5'	P_6'	P_7'	P_8'
Pancreatic Trypsin Inhibitor (Kunitz) Family[a]	...G	P	C	K	A	R	I	I	R	Y	F	Y...
		L		R	G	A	L	P	Q	F	Y	F
		R		Y	R	S	F	L	S	W	A	
		N		M	Q	Y	T	R	A	I	H	
				L		K	V	Q	L			
				-		F		S				
						H						
						M						
						P						
Pancreatic Secretory Trypsin Inhibitor (Kazal) Family[b]	...G	C	P	R	I	Y	N	P	V	C...		
	V		N	K	D	L	R	L	I			
	A		T	E	E	P	H	F	L			
	L		A	D	A	F	Q	R	H			
	D		M	A	L	H	S	E				
	M		L	L	Q	D	M					
	F		S	S	N	Q	K					
				M		M	D					
				V			I					
				Q			E					
				Y								
Streptomyces Subtilisin Inhibitor Family[c]	...M	C	P	M	V	Y	D	P	V	L...		
	A		T	K	Q	F				V		
Bowman-Birk Inhibitor Family[d]	...A	C	T	K	S	N	P	P	Q	C...		
	M		A	R		M		G	K			
	V			A		I		A	T			
	L			L		Q						
	I			F		Y						
	S			Y								

In general the P_1 reaction site residue corresponds to the specificity of its respective enzyme. Trypsin-like enzymes are inhibited by peptides with arginine or lysine in the reactive site; chymotrypsin-like proteases are inhibited by those with tyrosine, phenylalanine, leucine or methionine, while inhibitors with alanine or serine in the P_1 reactive site inhibit elastase-like enzymes (Laskowski *et al.,* 1971; Laskowski *et al.,* 1980). However, this is rule does not always apply, since inhibitors with P1 leucine or methionine were shown to be strong inhibitors of elastase in the Kazal family (Kato *et*

al., 1978) or alanine in the Bowman-Birk family (Ikenaka *et al.,* 1978). Other trypsin inhibitors with arginine or lysine in the P1 site also inhibit chymotrypsin (Rigbi, 1971).

Because the Kazal-like domain was present in the sequence of TFP4 protease inhibition assays were conducted to assess the protein's functionality. Trypsin, chymotrypsin and elastase were chosen, representing the three classes of serine proteases.

Trypsin will hydrolyze peptides on the C-terminal site of lysine or arginine residues. The rate of hydrolysis is slower if an acidic residue is on either side of the cleavage site and no cleavage occurs if proline is on the carbonyl side of the cleavage site (Walsh, 1970; Burdon *et al.,* 1989).

Trypsin inhibitors are found in a variety of sources such as bovine pancreas, human plasma, and soybean. The soybean inhibitor has been studied extensively; at 24000 Da molecular weight it is a single peptide chain cross linked by two disulfide bridges. A variety of synthetic compounds have been employed as trypsin substrates including N-benzoyl-DL-arginine-p-nitroanilide (Stewart, 1973), N-glutaryl-L-phenylalanine-p-nitroanilide or p-nitrophenyl-p-guanidinobenzoate (Ford *et al.,* 1973).

The substrate used in our assays was N-benzoyl-L-arginine-ethyl-ester (BAEE), an assay based on that of Laskowski (Laskowski, 1955). Trypsin hydrolysis of BAEE is monitored at 253 nm (figure 5.2). The assay was performed to obtain a known activity value for trypsin following the addition of a trypsin inhibitor measuring residual trypsin activity.

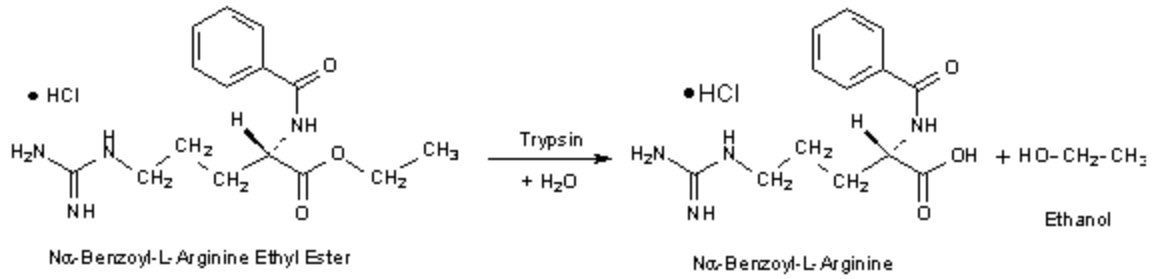


Figure 5.2 Hydrolysis of BAEE by trypsin

Chymotrypsin is a digestive system serine protease. It preferentially cleaves peptide bonds where the P1 is a tyrosine, tryptophan or phenylalanine. In vitro it can use substrate analogs such as benzoyl-L-tyrosine-ethyl-ester (BTEE) or N-acetyl-L-phenylalanine-p-nitrophenyl (Wilcox, 1970).

We used BTEE as substrate based on the method of Hummel (Hummel, 1959).Chymotrypsin hydrolyzes BTEE monitored by absorption at 256 nm (figure 5.3). We obtained an activity value for chymotrypsin followed with addition of an inhibitor allowing for measurement of residual enzymatic activity.

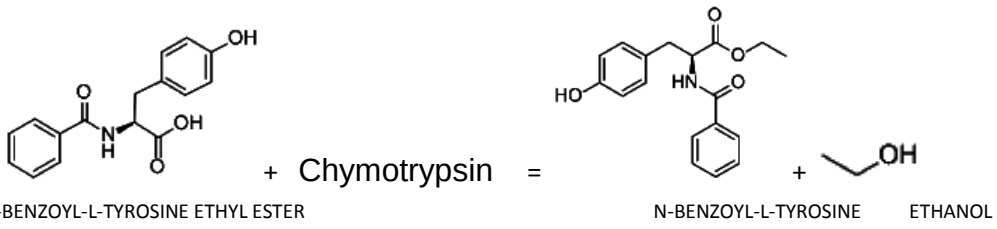
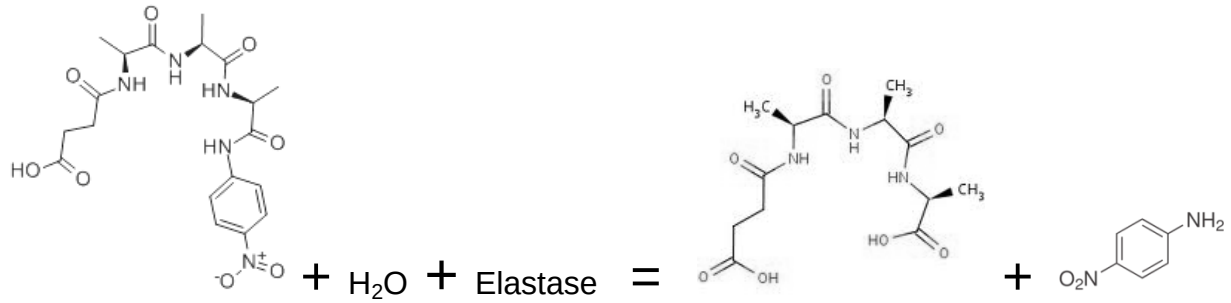


Figure 5.3 Hydrolysis of BTEE by chymotrypsin.

Elastases belong to families of serine proteases, cysteine proteases or metalloproteases and are capable of solubilizing fibrous elastin. Mammalian elastases occur mainly in pancreas, leukocytes and serum. Various synthetic compounds were

used as substrates in elastase assays such as succinyl-(l-alanine)$_3$-p-nitroanilide (Bieth *et al.,* 1974) or N-methoxysuccinyl-ala-ala-val-chloromethyl ketone (Stein, 1986).



N-Succinyl-Ala-Ala-Ala-*p*-nitroanilide N-Succinyl-Ala-Ala-Ala p-Nitroaniline

Figure 5.4 Hydrolysis of N- succinyl-(l-alanine)$_3$-p-nitroanilide by elastase

We used N- succinyl-(l-alanine)$_3$-p-nitroanilide (SucAla$_3$NA) based on the method of Feinstein (Feinstein *et al.,* 1973). Elastase hydrolyzes SucAla$_3$NA measuring light absorption at 410 nm (figure 5.4), performed to obtain a known activity value for elastase followed with addition of an elastase inhibitor, measuring residual enzymatic activity.

Materials and Methods

For trypsin, we mixed 750 µl of 0.1 mM BAEE, 50 µl of 6 ug/ml trypsin, 50 µl 0.067 sodium phosphate buffer pH 7.6 and 150 µl dH_2O for a total volume of 1.0 ml. The light absorbance was measured at 253 nm at room temperature using a Beckman DU 530 spectrophotometer.

For chymotrypsin, the final reaction mixture concentrations were 38 mM Tris-HCl, 0.55 mM BTEE, 23% (v/v) methanol, 53 mM $CaCl_2$, 1.7 µM HCl and 3.3 µg chymotrypsin in a total volume of 1 ml. The light absorbance was measured at 256 nm at room temperature using a Beckman DU 530 spectrophotometer.

For elastase the reagents were 100mM Tris buffer pH 8.0, 4 mM $SucAla_3NA$, and 6 µg/ml elastase. Both the enzyme and the substrate were diluted in Tris buffer and mixed together in a reaction with a total volume of 1 ml. The light absorbance was measured at 410 nm at room temperature using a Beckman DU 530 spectrophotometer.

Results and Conclusions

Trypsin Assays

Trypsin hydrolyses the ester bond of N-benzoyl-L-arginine-ethyl-ester (BAEE) resulting in N-benzoyl-L-arginine and ethanol. The plot of enzymatic activity vs. time is presented in figure 5.5.

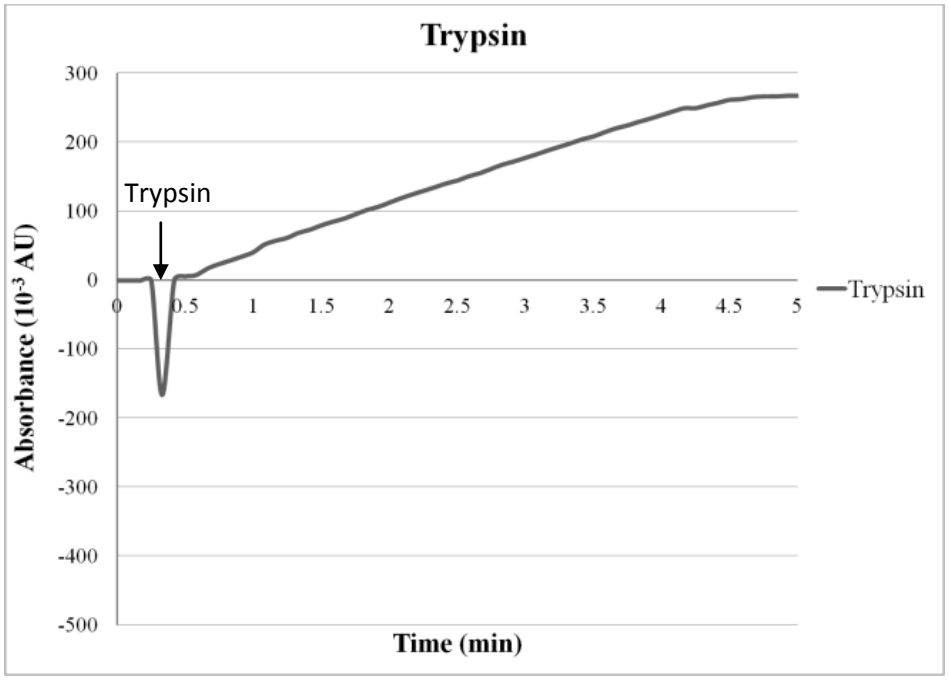


Figure 5.5 Trypsin activity using BAEE as substrate. Absorbance was measured at 253 nm and 25C.

Trypsin was used at 6 µg /ml with BAEE at 0.75 M in a total volume of 1 ml. The rate was $70x10^{-3}$ U/min with an R^2 coefficient of 0.9979. The R^2 coefficient of determination is a statistical measure of how well the regression line approximates the real data points

with values between 0 and 1. An R^2 of 1.0 indicates that the regression line perfectly fits the data.

The enzymatic activity of trypsin was inhibited by a protease inhibitor cocktail from Roche which is a mixture of serine, cysteine and metalloprotease inhibitors (figure 5.6).

According to Roche, 93% of trypsin (0.002 µg/ml) inhibition is achieved immediately and 73% of trypsin activity is inhibited after 60 minutes. In the first step trypsin was mixed with the BAEE substrate and after one minute the inhibitor was added.

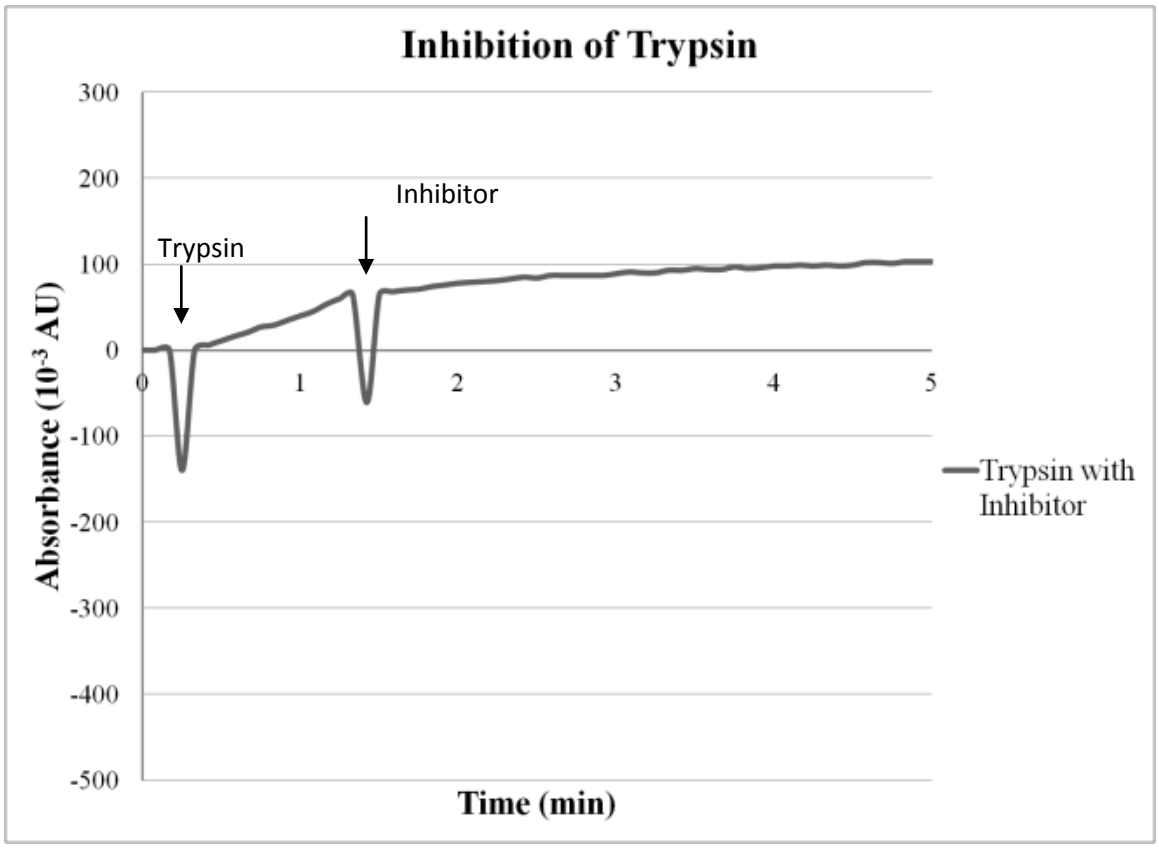


Figure 5.6 Inhibition of trypsin activity by Complete^R protease inhibitor cocktail (Roche)

The rate of the first reaction was 63.5×10^{-3} U/min with $R^2 = 0.994$ and the rate of the inhibited reaction was 12.5×10^{-3} U/min ($R^2 = 0.94$), which would be expected since not all trypsin was inhibited.

The assay with TFP4 was also done sequentially. Trypsin and BAEE were mixed together and data recorded for a minute, then TFP4 was added at 6 µg /ml.

According to figure 5.7 TFP4 appears to have no effect on trypsin's ability to hydrolyze BAEE. The rate for trypsin in the first phase was $67x10^{-3}$ U/min (R^2 = 0.997) and it changed to $82.4x10^{-3}$ U/min (R^2 = 0.9938). In fact, it seems the reaction accelerated after addition of the termite protein.

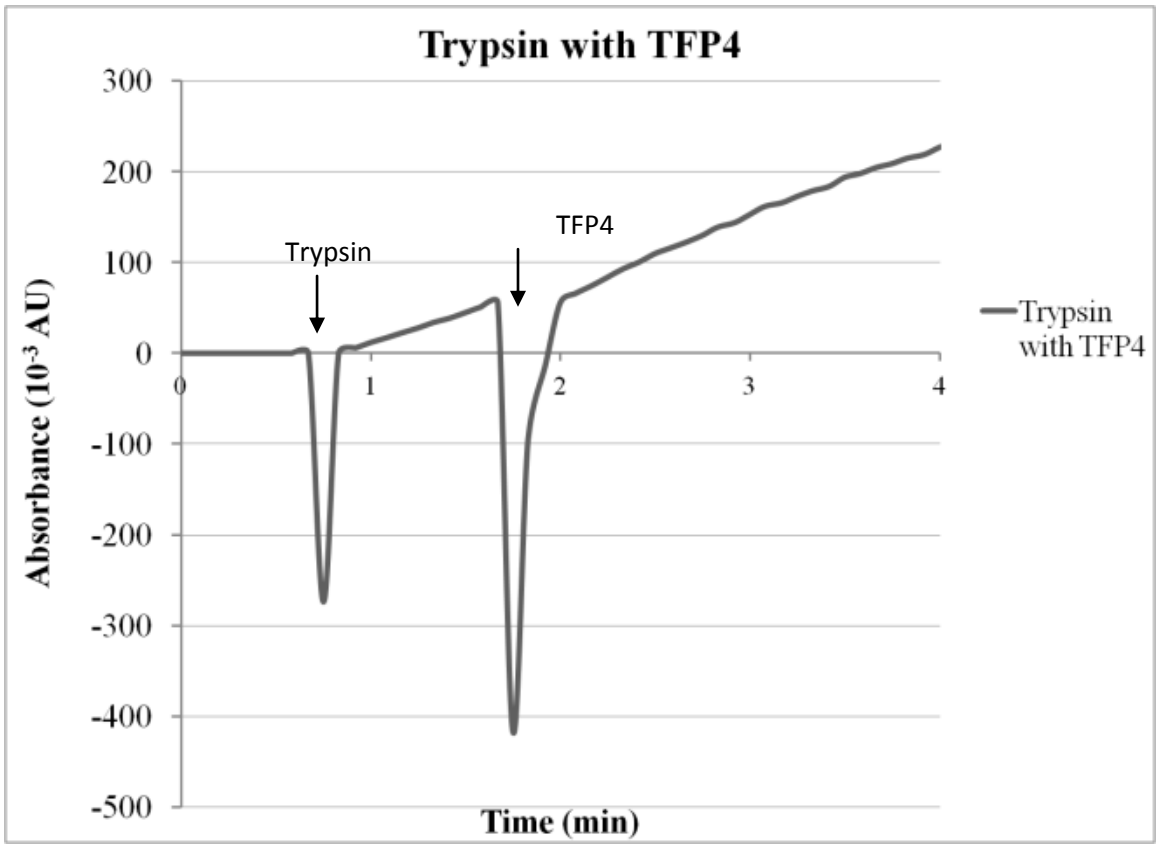


Figure 5.7 Trypsin activity measured before and after addition of TFP4.

Figure 5.8 presents a composite of trypsin's enzymatic activity and table 4 summarizes the important results. The three trypsin assays show remarkable consistency with similar rates of reaction, while adding TFP4 to trypsin seems to enhance enzyme's

activity. The inhibition reaction with the Roche cocktail has a significantly lower rate so the conclusion drawn is that termite protein TFP4 does not inhibit trypsin

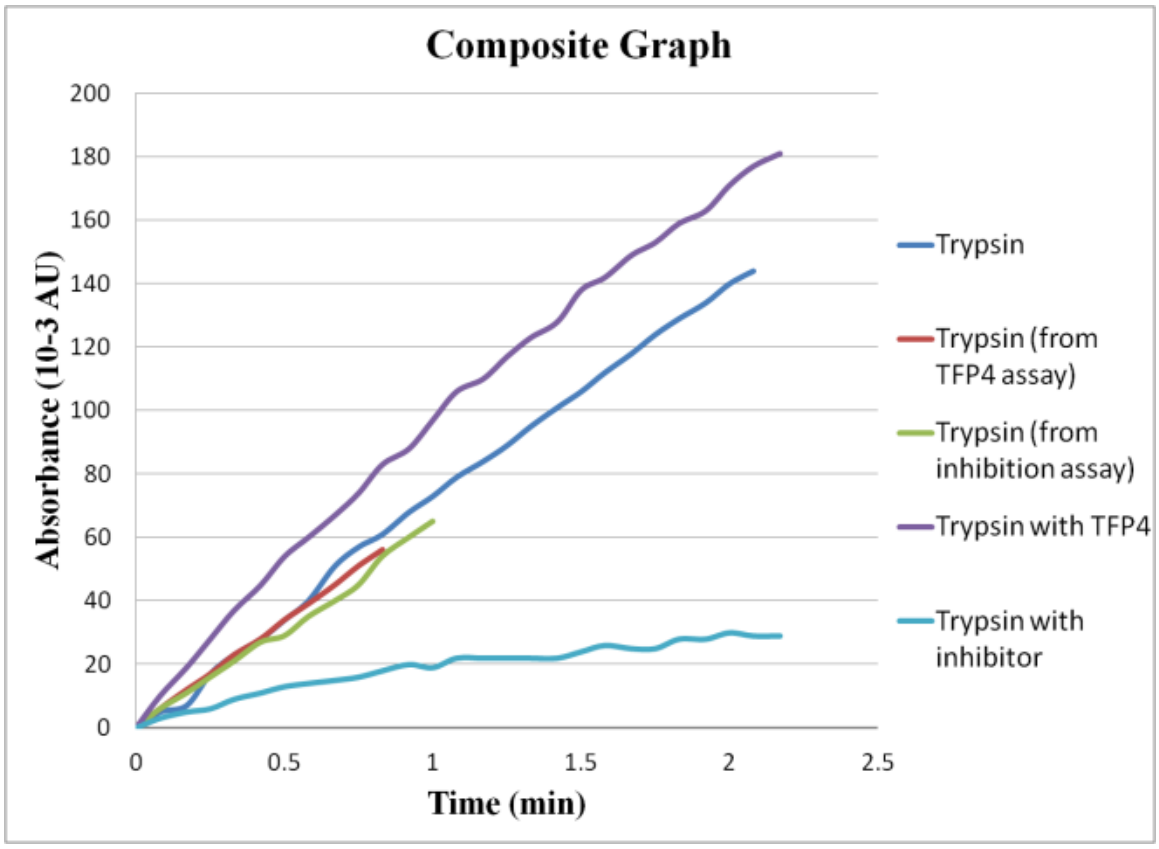


Figure 5.8 Composite graph of trypsin activity.

Table 5.4 Reaction rates and coefficients of determination for the trypsin assays.

Reaction	Rate (10^{-3} U/min)	R^2
Trypsin	70.642	0.9979
Trypsin (from TFP4 assay)	82.469	0.9938
Trypsin (from inhibition assay)	63.568	0.994
Trypsin with TFP4	67.026	0.9997
Trypsin with Inhibitory Cocktail	12.531	0.9384

Elastase Assays

Elastase assays were done using N- succinyl-(l-alanine)$_3$-*p*-nitroanilide as substrate. In the presence of water elastase hydrolyses the bond between the carbonyl carbon of the third alanine residue and the nitrogen of nitroaniline resulting in N- succinyl-alanine-

alanine-alanine and *p*-nitroaniline. The reaction (figure 5.9) contained 0.27mM substrate and 0.2 µg /ml of enzyme; it was monitored at 410 nm and 25^0C.

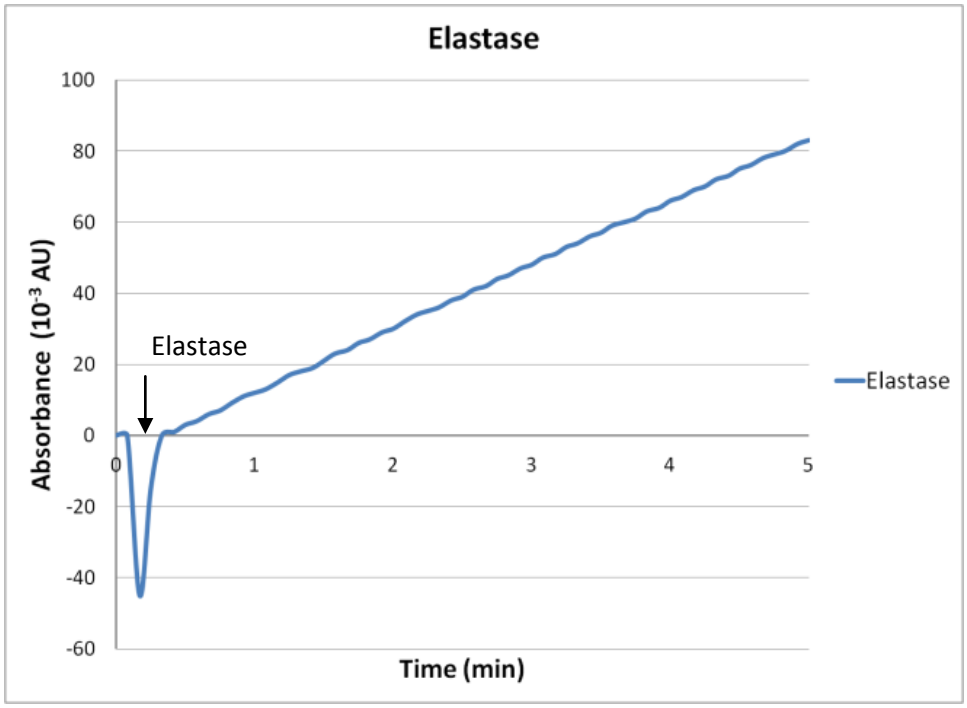


Figure 5.9 Elastase activity using N-succinyl-(l-alanine)$_3$-p-nitroanilide as a substrate. The reaction was monitored at 410 nm and 25^0C.

The rate of hydrolysis in this assay was 18.36 x 10^{-3} U/min with a "goodness of fit" coefficient R^2 = 0.9991.

When an inhibitor was added, the elastase's enzymatic activity was abolished (figure 5.10). The same inhibitor from Roche used in the trypsin assay was also used in this elastase assay at the same concentration.

Figure 5.10 Elastase inhibition using Complete^{R} protease inhibitor cocktail (Roche).

The rate of this reaction was 0 U/min with $R^2 = 1$.

In order to see what effect, if any, the termite protein TFP4 has on elastase, 0.2 µg /ml of TFP4 was added to a reaction of 0.27 mM substrate with 0.2 µg /ml elastase (figure 5.11).

For this reaction the observed rate was 1.15×10^{-3} U/min with $R^2 = 0.84$.

A final assay was performed using crude protein extract from the *Coptotermes formosanus* Shiraki frontal gland secretion (figure 5.12). The lyophilized extract was reconstituted in PBS buffer pH 7.0 and was used at a concentration of 1.2 µg /ml in the reaction. The rate of this reaction was 5.47×10^{-3} U/ml with $R^2 = 0.9957$. Although elastase activity is not totally abolished, it is slowed down significantly.

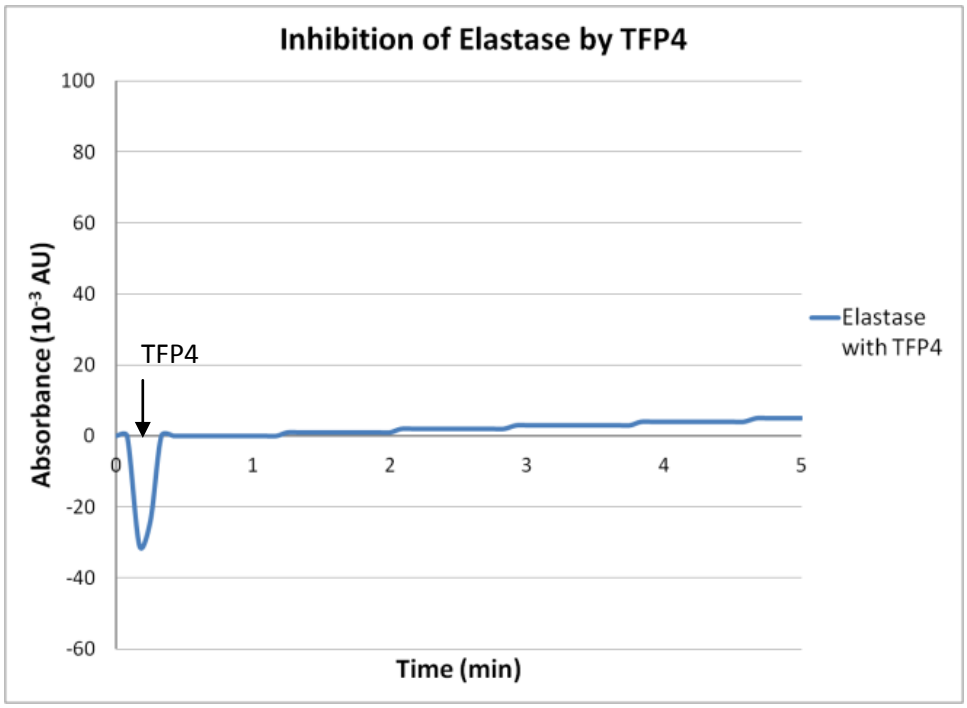


Figure 5.11 Inhibition of elastase by cloned soldier termite protein TFP4.

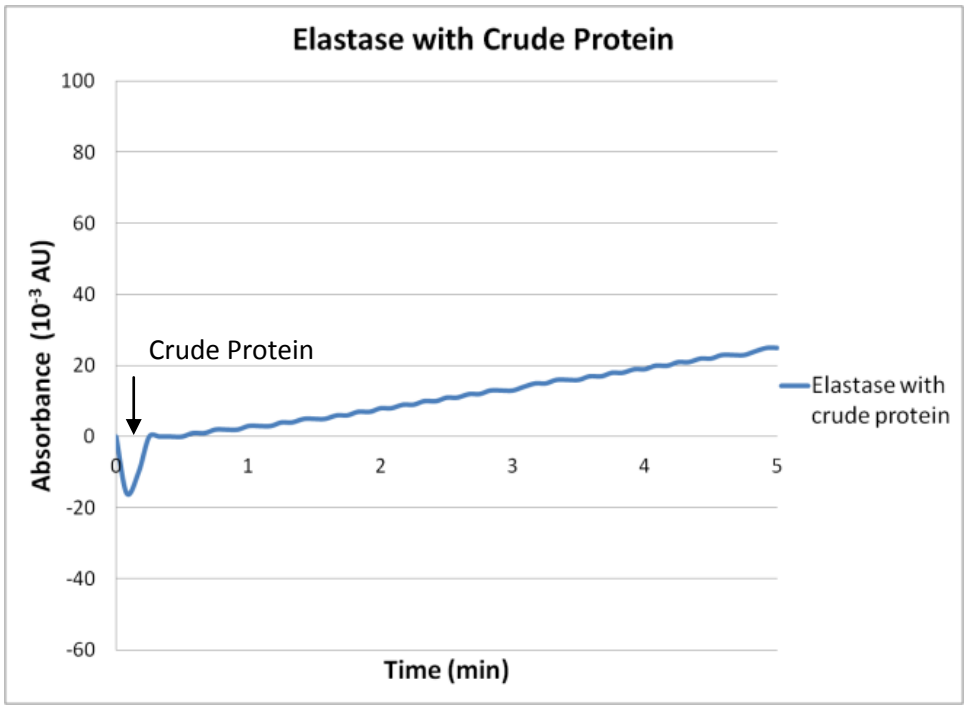


Figure 5.12 Partial inhibition of elastase enzymatic activity with an extract from the defense gland secretion of *Coptotermes formosanus* Shiraki.

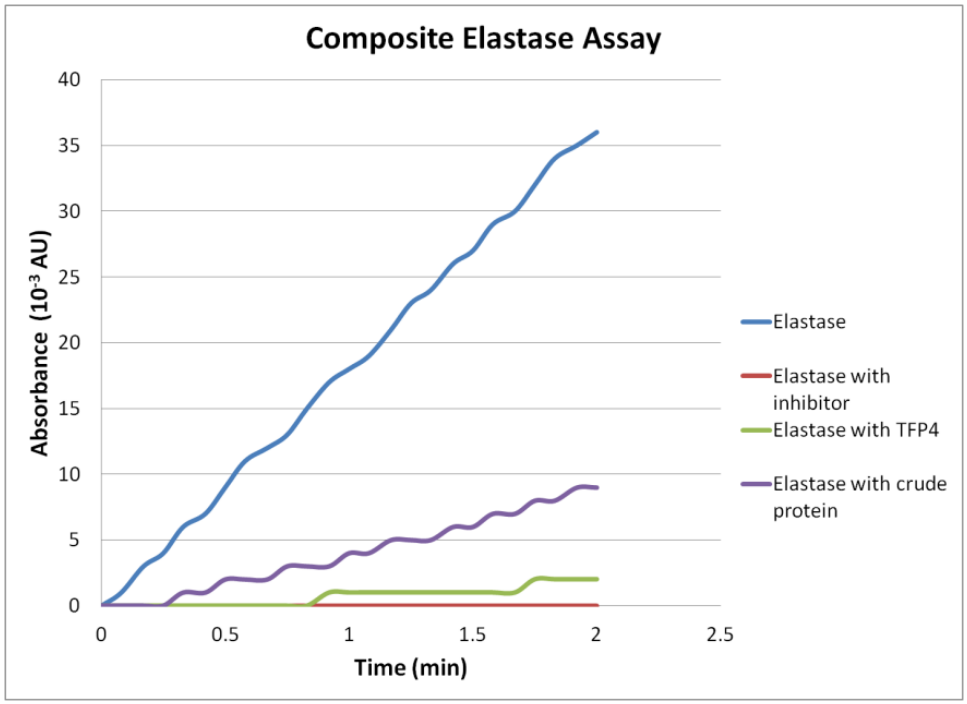


Figure 5.13 Composite graph of elastase activity.

A composite graph of elastase enzymatic activity (figure 5.13) and table 5 summarize the important results of the elastase assays.

Table 5.5 Reaction rates and coefficients of determination for the elastase assays.

Reaction	Rate (10^{-3} U/min)	R^2
Elastase	18.355	0.9991
Elastase with inhibitor cocktail	0	1
Elastase with TFP4	1.1503	0.9674
Elastase with crude protein	5.4699	0.9957

The hydrolyzing activity of elastase (18.36 x 10^{-3} U/min) was abolished by the Roche protease inhibitor cocktail and was inhibited (1.15 x 10^{-3} U/min) by the termite protein TFP4. In addition, the original termite protein expressed by the termite soldiers and present in the crude extract slowed the reaction significantly (5.47 x 10^{-3} U/min), supporting the conclusion that TFP4 is an inhibitor of elastase.

Chymotrypsin Assays

Chymotrypsin was used to hydrolyze N-Benzoyl-L-Tyrosine- Ethyl-Ester (BTEE) to N-Benzoyl-L-Tyrosine and ethanol in a reaction monitored at 253 nm and 25^0C. Various concentrations of substrate and enzyme were used and the results using 0.5 µM, 0.2 µM and 0.1 µM of BTEE are presented.

A typical reaction contained 467 µl substrate, 500 µl reaction buffer and 33 µl enzyme (figure 5.14). When needed, other reactants were added to the reaction replacing 33 µl of master mix of BTEE and buffer.

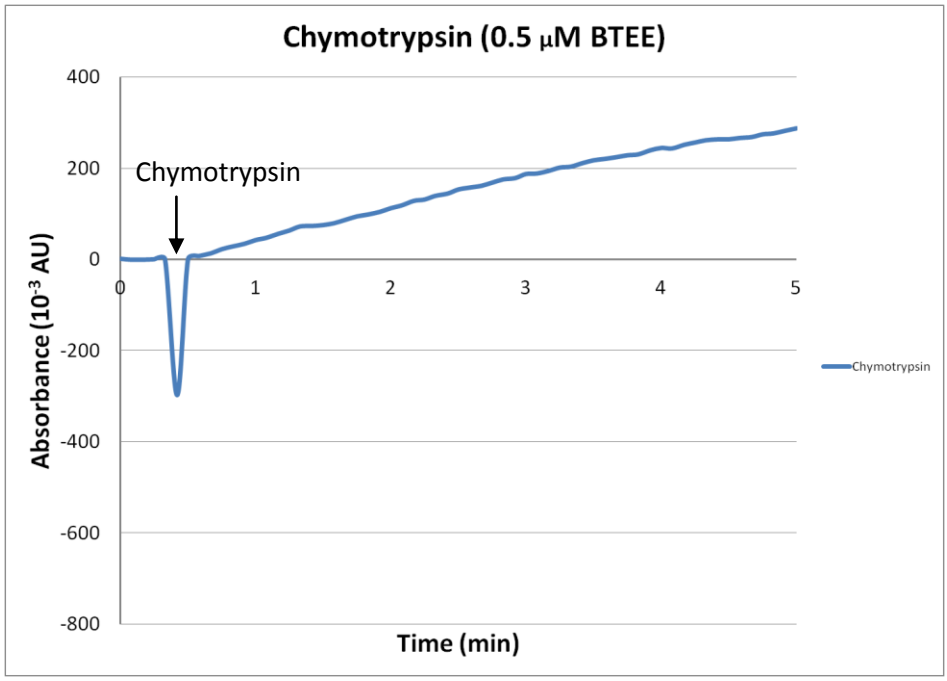
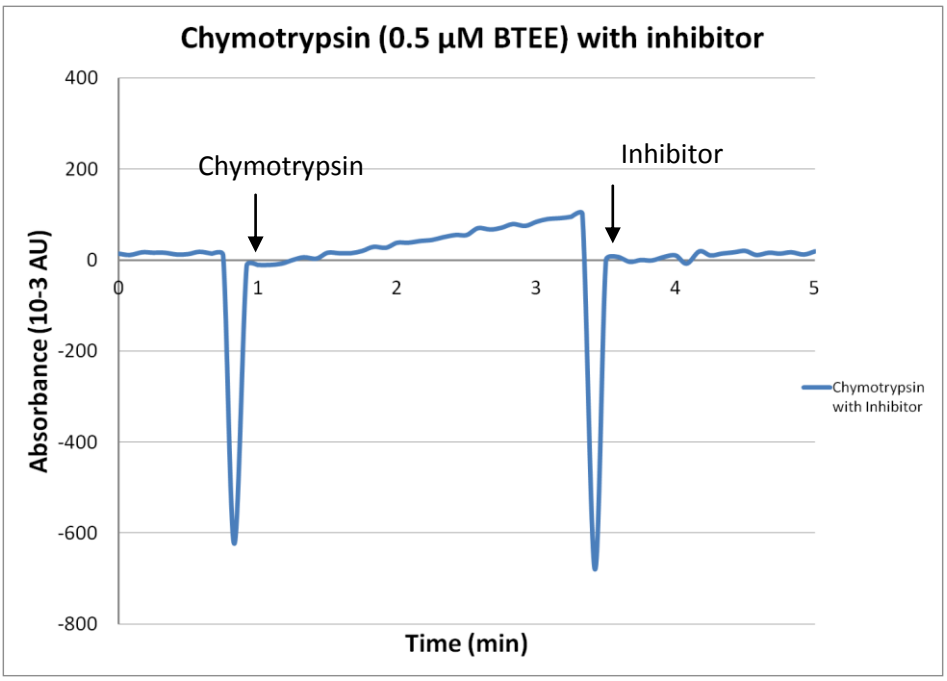


Figure 5.14 Chymotrypsin assay using 1 µg/ml enzyme and 0.5 µM BTEE.

a.



b.

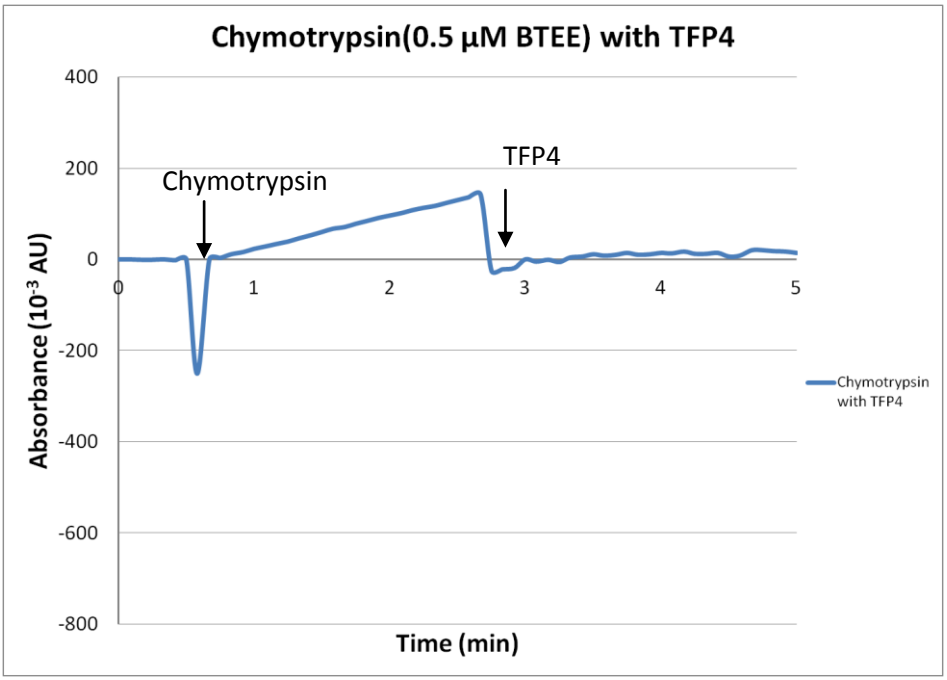


Figure 5.15 Chymotrypsin with inhibitor assay (a), Chymotrypsin with TFP4 assay (b).

For the reaction in figure 5.14, 33 µl of chymotrypsin were added for a total reaction concentration of 1 µg /ml. The substrate was hydrolyzed at a rate of 73.4×10^{-3}U/min with a confidence coefficient $R^2 = 0.9975$.

To assay the effect an inhibitor would have on chymotrypsin, 30 µl of a stock solution of 2 mg/ml Complete^{R} protease inhibitor cocktail from Roche were added to the reaction of chymotrypsin and BTEE (figure 5.15a). In the inhibition assay the reaction rate fell to 13.03 x 10-3 U/min with R^2 = 0.5521.

When TFP4 was added, the graph profile was similar to the protease inhibitor cocktail's assay (figure 5.15b). The reaction rate fell from 73.4x10^{-3} U/min to 10.8x10^{-3}U/min with R^2 equal to 0.5738. TFP4 was used at 6 µg /ml concentration and chymotrypsin at 1 µg/ ml. The low value for the confidence coefficient can be explained by the dispersion of data recorded following the addition of the reactants (inhibitor and TFP4).

Another assay was done using the crude protein extract from the defense gland secretion of *C. formosanus* termite soldiers (figure 5.16).

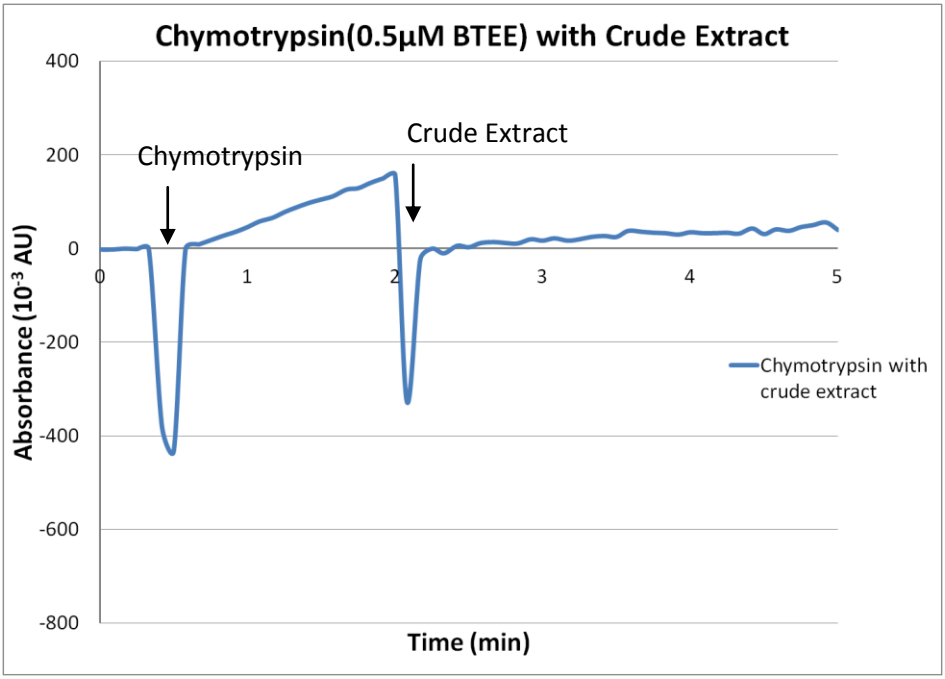
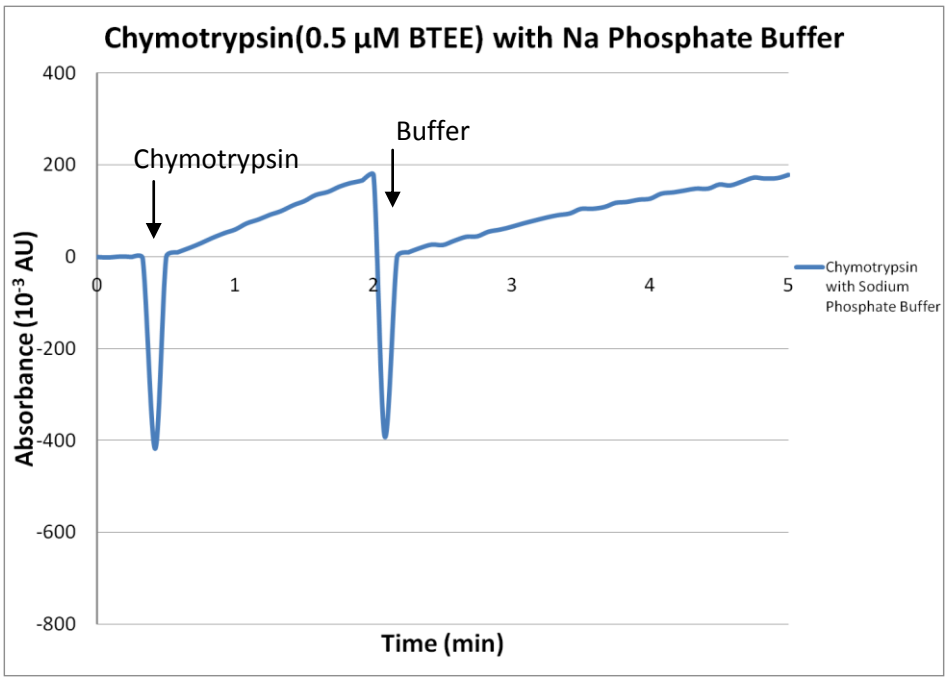


Figure 5.16 Chymotrypsin with crude protein extract from defense gland secretion of *C. formosanus*

a.



b.

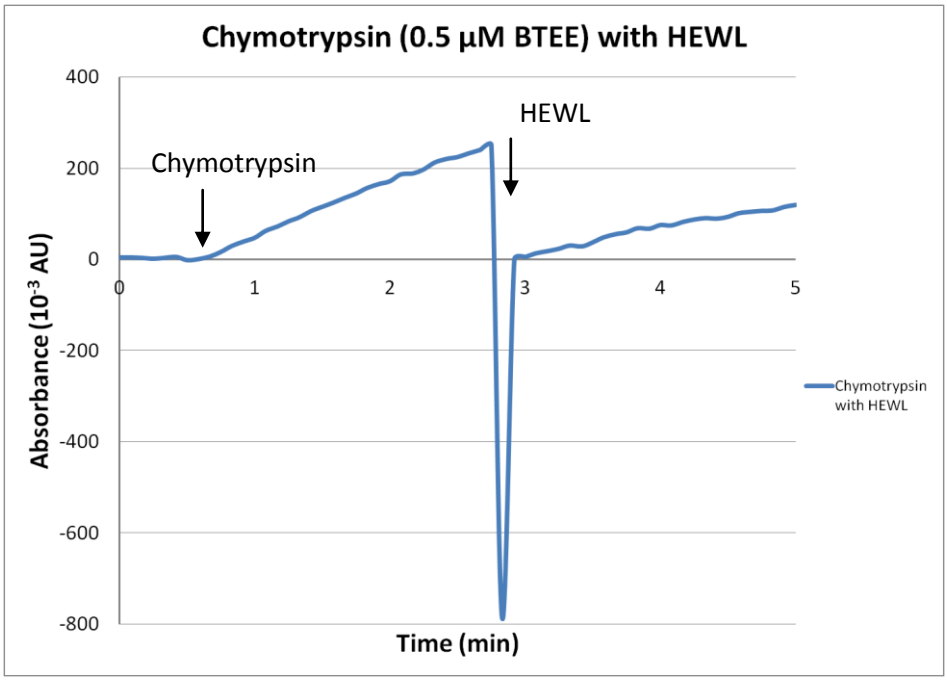


Figure 5.17 Chymotrypsin with sodium phosphate buffer assay (a); chymotrypsin with HEWL (b).

The protein concentration was 1.3 µg /ml and chymotrypsin activity was reduced after the addition of the extract to a rate of 16.66×10^{-3} U/min with $R^2 = 0.8626$. To exclude other factors which might influence chymotrypsin activity, an assay using the sodium

phosphate buffer (used for TFP4 refolding) was performed (figure 5.17a), along with an assay for HEWL, as a negative control (figure 5.17b).

In the sodium phosphate assay the reaction rate was 72.71×10^{-3} U/min with $R^2 = 0.9939$ and in the HEWL assay the hydrolysis of BTEE occurred at a rate of 70.26×10^{-3} U/min and $R^2 = 0.9982$. Both rates were similar to the control and TFP4 assays, while the confidence coefficient was over 0.99 indicating a high level of confidence in the data.

Other assays were conducted at 0.2 µM and 0.1 µM substrate concentrations. For both assays chymotrypsin was at 0.2 µg /ml and TFP4 at 1.9 µg /ml resulting in a ratio of 1:9.5 enzyme to TFP4. As reported above, in the 0.5 µM BTEE reactions the ratio was 1:6 enzyme to TFP4 (1 µg /ml chymotrypsin, 6 µg /ml TFP4).

The 0.2 µM assays for control and TFP4 are presented in figures 5.18 and 5.19.

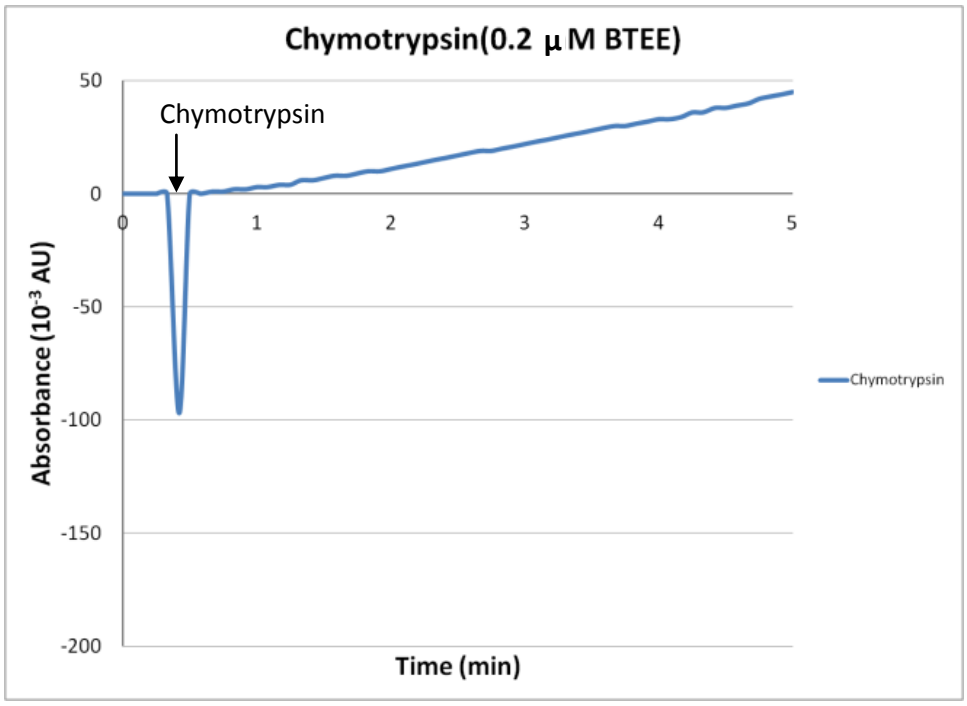


Figure 5.18 Chymotrypsin assay at 0.2 µM BTEE

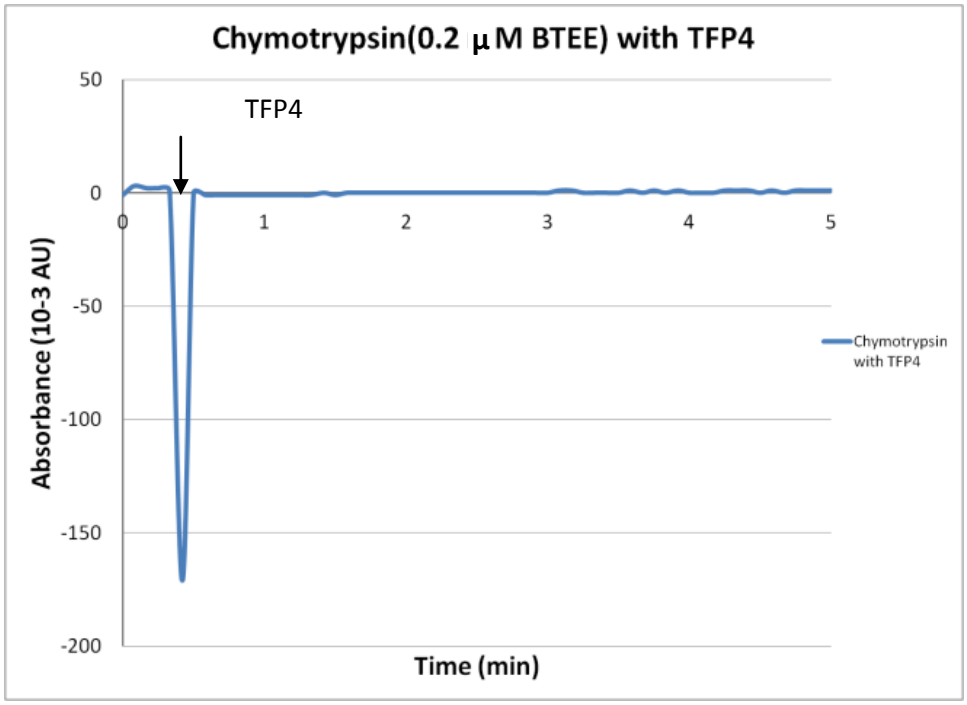


Figure 5.19 Chymotrypsin with TFP4 assay at 0.2 µM BTEE.

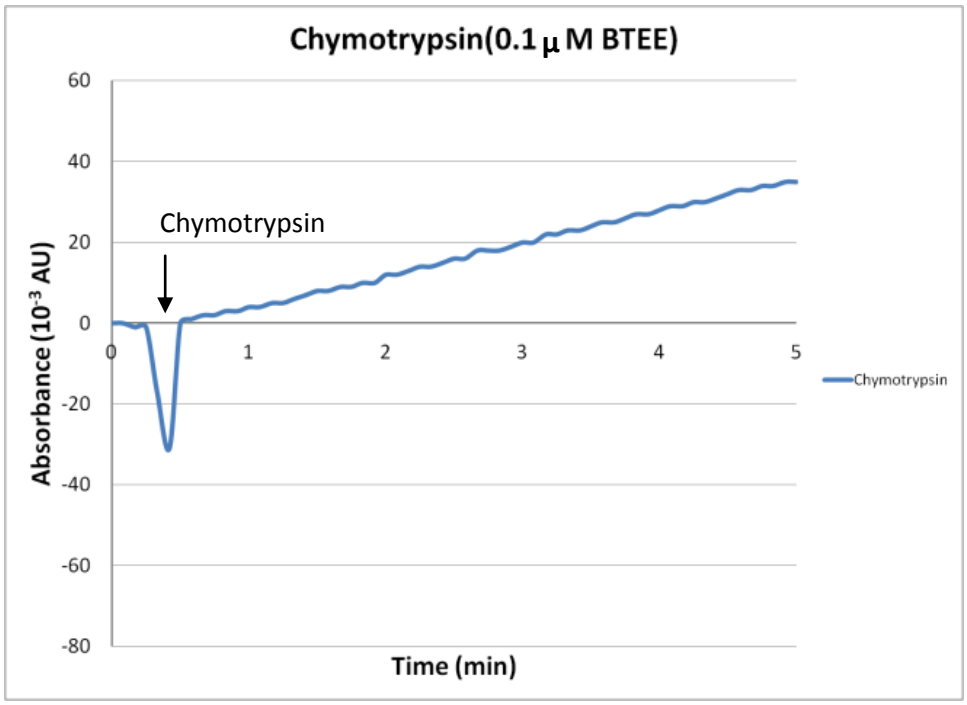


Figure 5.20 Chymotrypsin assay at 0.1 µM BTEE.

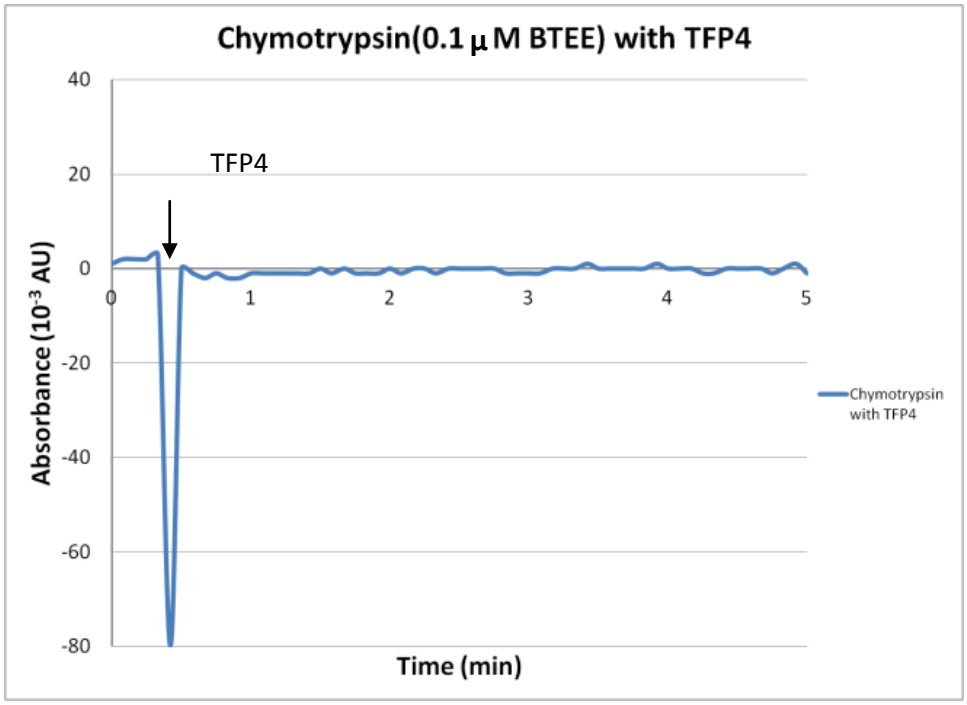


Figure 5.21 Chymotrypsin TFP4 assay at 0.1 µM BTEE.

The results of the chymotrypsin assays are summarized in figure 5.22 and table 6.

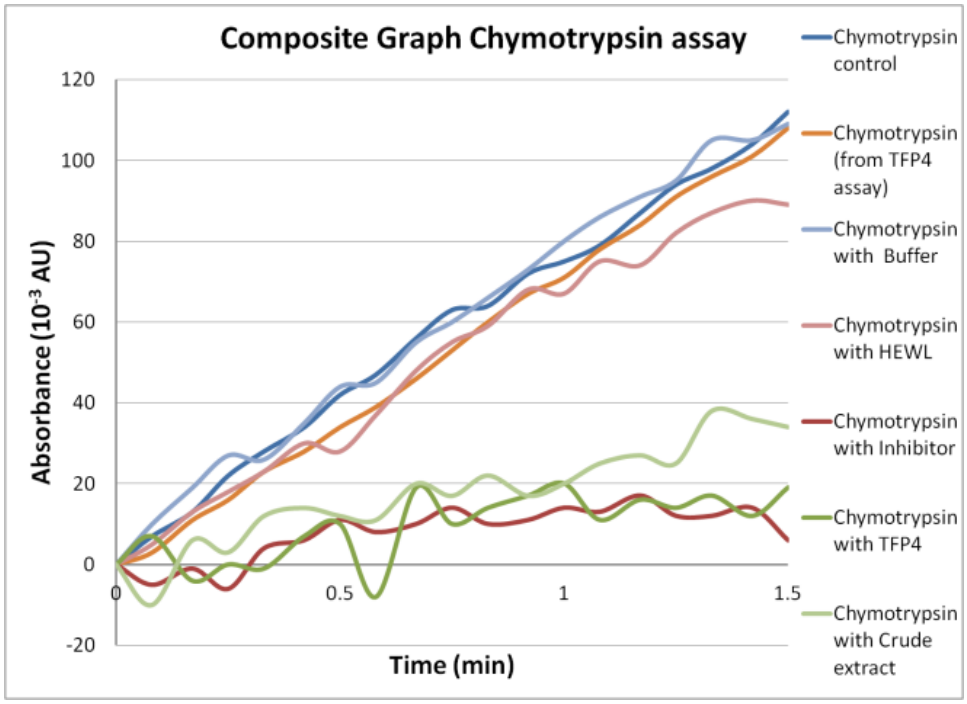


Figure 5.22 Composite graph of chymotrypsin assays.

The control reaction rate was 7.66×10^{-3}U/min while chymotrypsin and TFP4 reaction rate fell to 0.6739×10^{-3}U/min. The same reaction profiles were observed in assays done

with 0.1 µM BTEE, the hydrolysis rate for chymotrypsin alone (figure 5.5a) was 7.30 x10^{-3}U/min while the reaction rate after TFP4 was added was 0.50x10^{-3} U/min.

Table 5.6 Reaction rates and determination coefficient R^2 for the chymotrypsin assays

Reaction	Rate (10^{-3} U/min)	R^2
Chymotrypsin (0.5 mM BTEE)	73.747	0.9975
Chymotrypsin (0.5 mM BTEE from TFP4 assay)	73.378	0.9992
Chymotrypsin with Na Phosphate (0.5 mM BTEE)	72.707	0.9939
Chymotrypsin with HEWL (0.5 mM BTEE)	70.259	0.9882
Chymotrypsin with Inhibitor (0.5 mM BTEE)	13.028	0.5221
Chymotrypsin with TFP4 (0.5 mM BTEE)	10.803	0.5738
Chymotrypsin with Crude Extract	16.656	0.8626
Chymotrypsin (0.2 mM BTEE)	7.6581	0.9845
Chymotrypsin with TFP4 (0.2 mM BTEE)	0.6739	0.3885
Chymotrypsin (0.1 mM BTEE)	7.3022	0.9867
Chymotrypsin with TFP4 (0.1 mM BTEE)	0.5048	0.1453

Chymotrypsin hydrolyses N-Benzoyl-L-Tyrosine-Ethyl-Ester being significantly slowed with addition of TFP4 in all assays performed with 0.5 µM, 0.2 µM or 0.1 µM substrate and also by the crude protein extract. Therefore, the conclusion drawn from these experiments is that the termite soldier defense gland secretion protein TFP4 is an inhibitor of chymotrypsin.

An enzyme inhibitor binds to enzymes and decreases or abolishes their enzymatic activity. Binding of an inhibitor can prevent a substrate from entering the active site or impede the enzyme from catalyzing the reaction.

Two types of inhibitors were described – reversible and irreversible inhibitors. Reversible inhibitors bind the enzyme non-covalently through hydrophobic interactions,

hydrogen bonds or ionic bonds while irreversible inhibitors form covalent bonds with the enzyme and the inhibition cannot be reversed (Zubay, 1983).

There are four types of reversible inhibitions – competitive inhibition, uncompetitive inhibition, mixed inhibition and non-competitive inhibition.

In competitive inhibition the inhibitor has an affinity for the enzyme's active site and competes with a substrate for access to the catalytic pocket.

In uncompetitive inhibition the inhibitor binds to the enzyme-substrate complex and prolongs the time the complex stays formed, affecting the enzymes turnover values. In mixed inhibition the inhibitor binds to an allosteric site, a site different from the enzyme's active site. This event changes the enzyme's conformation (allosteric effect) and substrate binding is hindered.

Non-competitive inhibition is a special case of mixed inhibition where the inhibitor has equal affinity for the enzyme or the enzyme-substrate complex. The enzyme binds both the substrate and the inhibitor but the complex cannot form a product, it can only reverse to enzyme-substrate or enzyme-inhibitor complex (Metzler, 2001).

By observing the enzyme activity under various substrate and inhibitor concentrations, a Lineweaver-Burke plot can be constructed and the type of inhibition mechanism determined (Leatherbarrow, 1990).

Following the assays which established that TFP4 is an inhibitor of chymotrypsin, more assays were performed for the construction of a Lineweaver-Burke plot. The enzyme concentration was 0.2 ug/ml for all reactions while the amount of substrate and TFP4 varied according to Table 7.

The data was processed using Kinetics software and a plot generated (figure 5.23)

Table 5.7 Substrate and TFP4 concentrations used to determine reaction rates for the Lineweaver-Burke plot.

BTEE concentration	TFP4 concentration	V (U/min)	1/V
0.05 µM	0 µg /ml	0.0011	909
0.05 µM	1 µg /ml	0.0007	1430
0.05 µM	2 µg /ml	0.0003	3330
0.05 µM	4 µg /ml	0	n/a
0.075 µM	0 µg /ml	0.0013	769
0.075 µM	1 µg /ml	0.001	1000
0.075 µM	2 µg /ml	0.0005	2000
0.075 µM	4 µg /ml	0	n/a
0.1 µM	0 µg /ml	0.002	500
0.1 µM	1 µg /ml	0.0012	833
0.1 µM	2 µg /ml	0.0004	2500
0.1 µM	4 µg /ml	0.0004	2500
0.2 µM	0 µg /ml	0.0028	357
0.2 µM	1 µg /ml	0.0019	526
0.2 µM	2 µg /ml	0.0012	833
0.2 µM	4 µg /ml	0.0006	1670
0.3 µM	0 µg /ml	0.0041	244
0.3 µM	1 µg /ml	0.0028	357
0.3 µM	2 µg /ml	0.0014	714
0.3 µM	4 µg /ml	0.001	1000

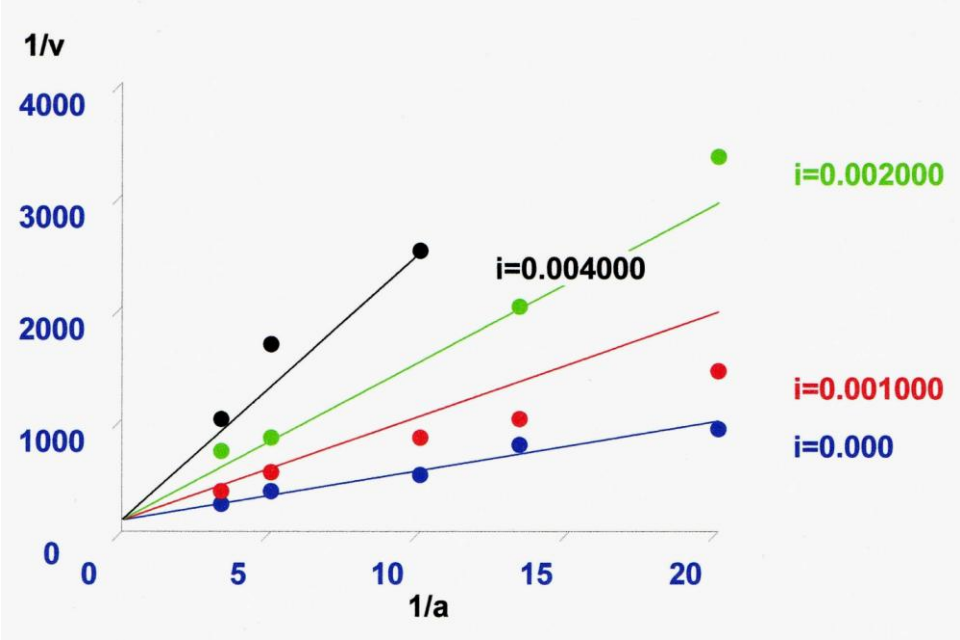


Figure 5.23 Lineweaver-Burke plot for chymotrypsin inhibition by TFP4.
This plot is representative for competitive inhibition with the formula:

$v = (V_m \times [S]) / (K_m \times (1 + [I]/K_i') + [S])$, where [S] and [I] are BTEE and TFP4 concentrations.

In conclusion, the chymotrypsin assays showed the termite soldier defense secretion protein TFP4 to be an inhibitor of chymotrypsin using a competitive inhibition mechanism.

CHAPTER SIX: NMR STUDIES

Introduction

NMR spectroscopy is a technique capable of determining the structures of biological molecules like proteins and nucleic acids at atomic resolutions. It measures the absorbance of energy due to changes in nuclear spin orientation.

NMR active atoms are any atoms with an odd number of neutrons and/or an odd number of protons. The most frequently used are ^{1}H (1 proton, 0 neutrons), ^{13}C (6 protons, 7neutrons) and ^{15}N (7 protons, 8 neutrons). All these isotopes have a characteristic spin of I=1/2. In theory a spinning charge generates a magnetic field with a magnetic moment proportional to the spin. If an external magnetic field is applied, the atom's magnetic field aligns with this external field and two spin states exist: m = 1/2, low energy, parallel to the external magnetic field and m = -1/2, high energy, antiparallel to the external magnetic field. The energy difference between these two spin states is proportional to the intensity of the applied magnetic field. Irradiation with radio frequency (rf) energy will cause the excitation of the nuclei from the lower state to the higher state, when the nuclei become resonant (Kemp, 1991).

Proton nuclei in different compounds will behave differently when the rf energy is applied. The electrons are charged particles and they respond to the external magnetic field by generating a secondary magnetic field which shields the nucleus. The consequence is the magnetic field applied must be increased for the resonance to occur. As a result, NMR resonance signals are dependent on both the intensity of the external field and the rf energy applied. To compare the different signals tetramethylsilane is used as a reference compound – $(CH_3)_4Si$ or TMS. The difference

between the resonant reference frequency and the resonant sample frequency, divided by the rf used is called the chemical shift, a number that reflects the difference between the reference and the sample in ppm. By monitoring the emission of absorbed energy, an NMR spectrum is obtained from a sample, with peaks at different chemical shifts reflecting the resonance of each proton. The structure can be determined by using tables with chemical shifts for known compounds (Kemp, 1991).

^{1}H NMR spectra exhibit chemical shifts, splitting patterns and different peak intensities. Chemical shifts arise when the proton is affected by the electronegativity of neighboring atoms or groups. Splitting patterns, referred as spin-spin interactions, are generated by the effect of adjacent hydrogens on the nuclei of interest. How many adjacent protons exist is given by the p=n+1 rule where p is the number of peaks, and n is the number of adjacent protons. So, for a triplet, p=3 and n=2 adjacent protons. Peak intensities are in direct correlation with how many protons resonate at the same frequency, i.e. how many hydrogens bind the same carbon (Silverstein *et al.,* 2005).

For the NMR of biological molecules, using ^{1}H NMR is not enough to solve the 3D structure because of all the interactions inside the molecule. The peaks overlap each other, it is difficult to extract peak intensity information, and peaks are too broad so the fine detail is lost. The solution is to repeat the NMR using another isotope, like^{13}C or ^{15}N and combine the data with ^{1}H NMR. Heteronuclear Signal Quantum Coherence (HSQC) experiments record the chemical shift of ^{1}H directly while the chemical shift of the second nucleus (^{15}N or ^{13}C) is recorded in the indirect dimension.

When the protein sequence is known, as in this study, it is important to get a resonance assignment for it, to determine which chemical shift corresponds to which atom.

Different techniques were developed which allow the generation of data that is used in assigning interatomic distances and torsion angles (Balei, 2005).

The techniques usually used are homonuclear correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY). COSY and TOCSY transfer magnetization through the chemical bonds between adjacent protons. NOESY transfers magnetization through space, it will show peaks for all the protons that are close in space, regardless if they are in the same spin system or not.

Homonuclear correlation spectroscopy (COSY), for example, consists of an rf pulse, followed by evolution time, followed by a second rf pulse and a measurement time. The peaks recorded result from magnetization transfer. In COSY the transfer occurs through the bonds and interresidual coupling is observed. Intraresidual coupling is observed through TOCSY. This technique creates correlations between all protons within a spin system, which usually is an amino acid residue. When ^{15}N is used, only the signals from the protons which interact with ^{15}N bound protons are detected, creating an additional layer for the interpretation of the results. Polarization transfer between nuclear spins – the nuclear Overhauser effect (NOE) – is observed through space not through bonds so the distance between atoms can be calculated using NOESY (Silverstein et al., 2005).

From the interpretation of the results generated by these techniques two types of restraints are generated – distance restraints and angle restraints. The data is fed to a computer and multiple NMR structures are generated (an ensemble), reflecting the fact that there are multiple solutions to the set of restraints provided. The structures will converge if the data is sufficient to dictate a specific fold for the protein (Balei, 2005).

Materials and Methods

^{15}N labeled TFP4 protein expression

E. coli strain BL21(DE3) cells were transformed with pET46TFP4 plasmid and were grown in M9 minimal media (13 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 1 g/L $^{15}NH_4Cl$, 2mM $MgSO_4$, 0.1 mM $CaCl_2$) at 37C for 16 hours. The culture was diluted 1:10 (v/v) with M9 minimal media, incubated 30 minutes followed by addition of 0.1 mM IPTG. After 4 hours of incubation at 37^0C the cells were centrifuged 10 minutes at 5000 rpm and the ^{15}N labeled protein extracted and purified on a Ni column following the protocol previously used. Additional purification using a G-100 column was performed.

NMR spectroscopy was done using the LSU facility.

Protein structure prediction used the I-Tasser Suite developed by the Zhang Lab at the University of Michigan.

Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco (supported by NIH P41 RR001081).

Results and Conclusions

Figures 6.1 and 6.2 represent the results from the NOESY scan and the HSQC experiment on the ^{15}N labeled TFP4 protein.

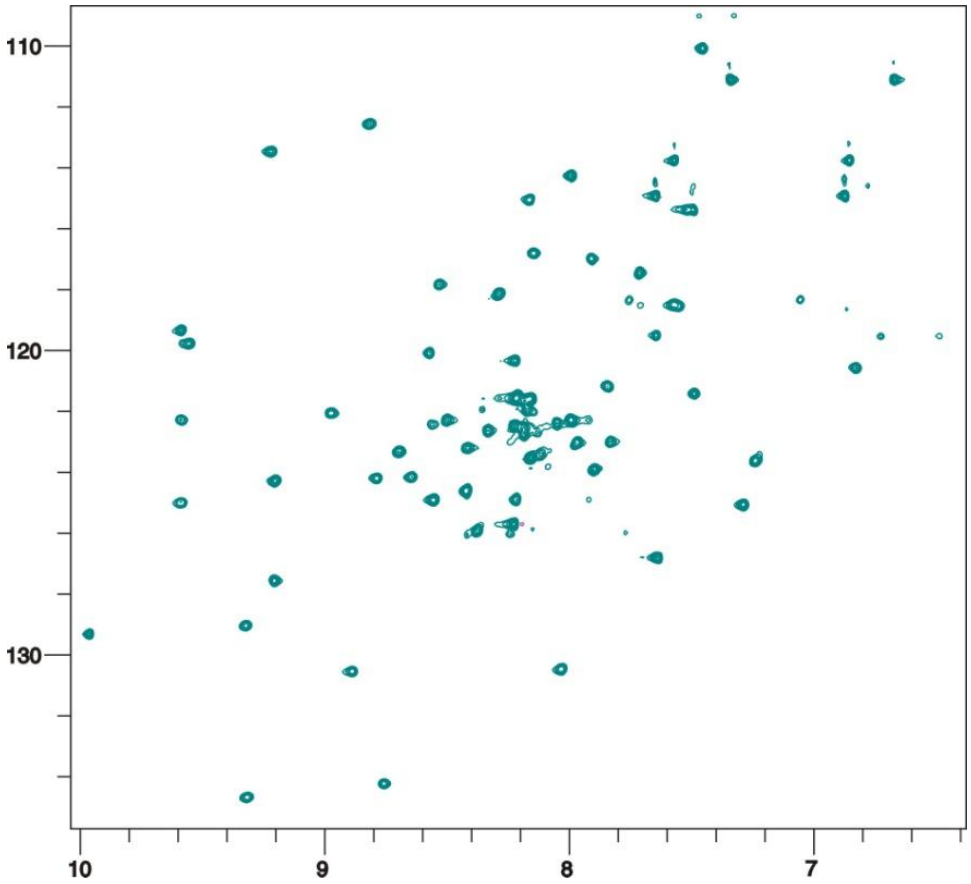


Figure 6.1 Nuclear Overhauser Effect Spectroscopy (NOESY) scan of ^{15}N labeled TFP4 protein. The signals are spread apart indicating a folded protein.

The method of protein structure determination using NMR spectra is, more or less, an all-or-none event. Data is accumulated and checked for consistency with known parameters derived from the protein's primary structure (number and type of amino acid residues) and is introduced into a computer program for analysis and structure prediction.

The NOESY plot provides information about the number of residues, folding, degradation and purity of the protein. Proline gives no signal and cannot be detected while the N-terminal His-tag sequence comprises a flexible structure and does not appear in the picture, as it should be expected for a sequence that was added when the protein was engineered.

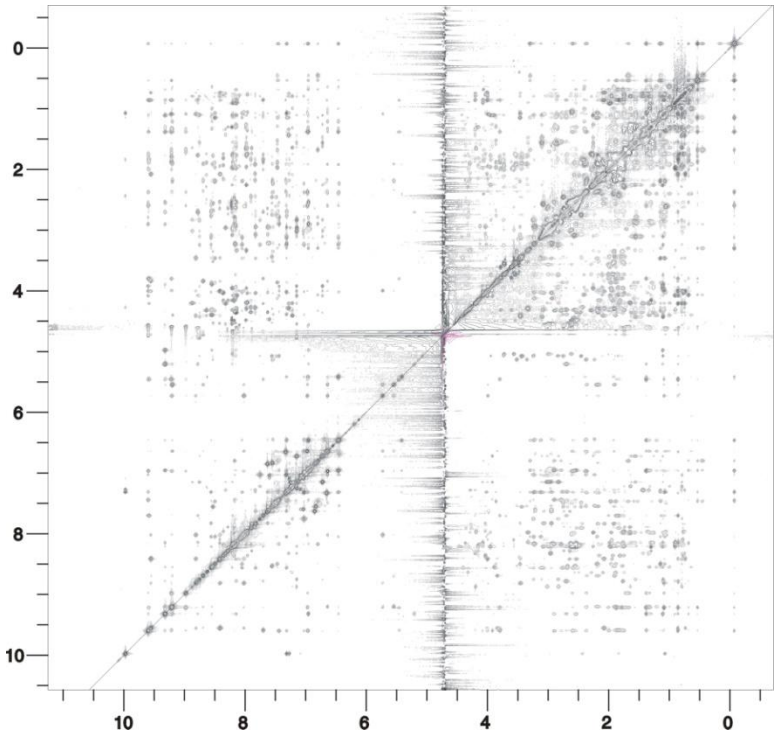


Figure 6.2 Heteronuclear Signal Quantum Coherence (HSQC) plot of ^{15}N labeled TFP4 protein

All the other amino acid residues, 55 in all, are accounted for and give clear signals, a sign of good concentration and high purity of the sample. A denatured, linear protein would give off signals concentrated in the middle of the plot, but here the signals are well dispersed indicating the protein is folded.

A protein structural model was generated by using I-Tasser Suite of computer programs developed by the Zhang Lab at the University of Michigan (figure 6.3). Various iterations were performed and the results were similar, leading to higher confidence in the features of the model chosen.

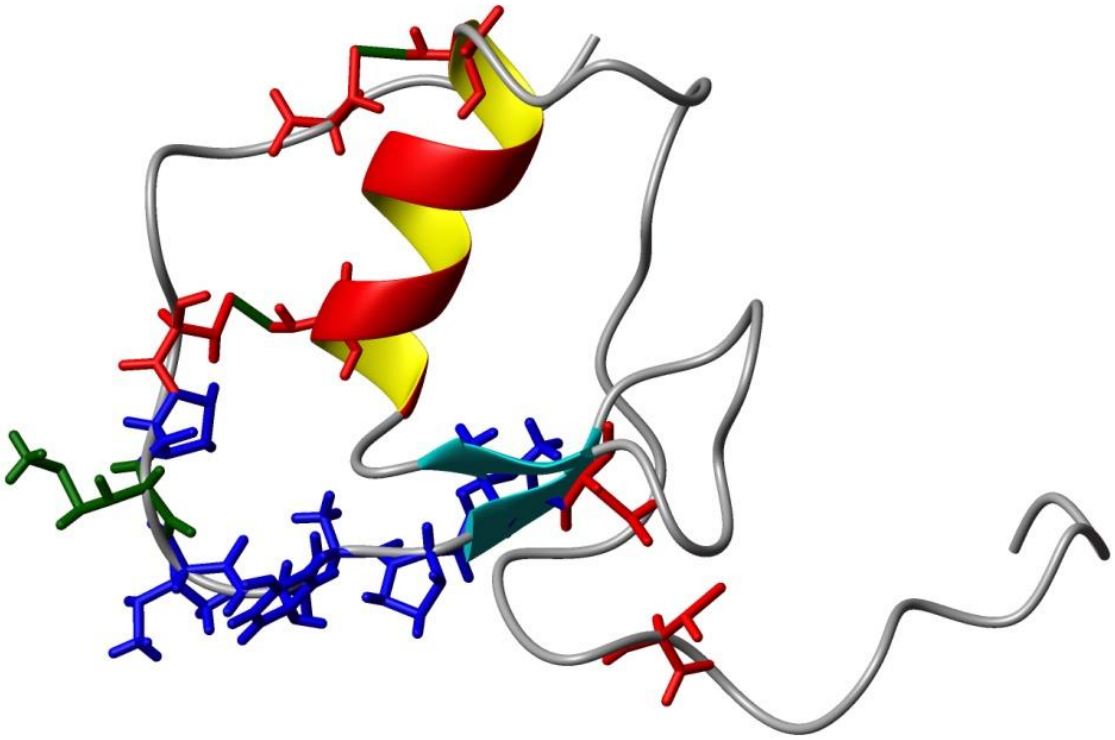


Figure 6.3 Model of TFp4, a protein in the defensive secretion from the frontal gland of *Coptotermes formosanus* Shiraki soldiers. Two β-strands, a potential β-hairpin and an α-helix were assigned to the structure.

The pattern of disulfide bonds in the known Kazal inhibitors is Cys1-Cys5, Cys2-Cys4 and Cys3-Cys6 (Laskowski and Kato, 1980), the same pattern being observed in this model. There is also a singular α-helix present, the same way a Kazal inhibitor structure is depicted.

In conclusion, the model presented is a good starting point for discussion of its different features and it will be refined with data from the NMR spectroscopy and X-ray diffraction which are ongoing.

CHAPTER SEVEN: CONCLUSIONS

Summary and Discussion

In this study the function and structure of a protein from the frontal gland secretion of soldiers of *Coptotermes formosanus* Shiraki were determined.

Starting with the observation that proteins were present in the secretion, the plan was to determine the N-terminal sequence and clone the corresponding cDNA into a vector for protein expression. The protein cloned was given the name TFP4 and the analysis of its sequence revealed that the protein's theoretical molecular weight was 6853 Daltons with a pI of 4.31 and -6.27 charge at pH 7.0. The sequence obtained was compared with the sequences of known proteins in NCBI's data base and similarities with other proteins analyzed.

Two functions were assigned to TFP4 - lysozyme and serine protease inhibitor.

Lysozyme Function

Lysozyme hydrolyzes b-(1,4)-N-acetyl-D-glucosamine-N-acetyl-muramic acid glycosidic linkages in cell walls of Gram positive bacteria. Two methods were used to assay this function - a zymogram assay and a fluorometric assay.

The zymogram assay showed clear zones of bacterial cell wall lysis where TFP4 has migrated on the gel, the same pattern that was observed in the zymogram with chicken egg white lysozyme. Bovine serum albumin (BSA) was loaded on the same gel but no clear zones were detected, consistent with the fact that BSA doesn't have lysozyme activity.

The fluorescent assay used *Micrococcus leisodeictikus* cells labeled with a fluorescent dye, fluorescein. If lysozyme activity is present the dye is released in solution and can

be detected by a spectrofluorometer. Fluorescence was observed in the samples with TFP4 directly proportional with the amount of protein used.

Based on data collected from two assay used for determination of lysozyme activity, the conclusions drawn was that TFP4, the protein from the frontal gland secretion of *Coptotermes formosanus* Shiraki soldiers has lysozyme activity and TFP4 is a lysozyme.

The classical lysozyme catalytic mechanism involves a catalytic acid, usually glutamic acid and a general base, usually aspartic acid. A study by Wohlkonig (Wohlkonig *et al.,* 2010) explored the structural relationship among lysozyme families in search of a common motif among members of the same family.

There are two elements which are common to proteins of lysozyme families: the helix of the enzymatic domain contains the catalytic glutamic residue and a β-hairpin structure has the highest amino acid conservation in a given family. The β-hairpin almost invariably contains the sequence GXXQ. Also a serine or threonine needs to be near the catalytic site in order to interact with the catalytic water molecule.

The structure of TFP4 satisfies only partially these requirements. The sequence GVSQ is present near the α-helix but there are no glutamic acid residues in the helix, only an aspartic acid, Asp32. However, there is a glutamic acid residue spatially close to the hairpin, Glu51. It sits close to Asp53, Asp 54 and Asp19. The β-hairpin sits in a pocket formed at the C-terminus of TFP4 and, theoretically, might participate in the catalytic

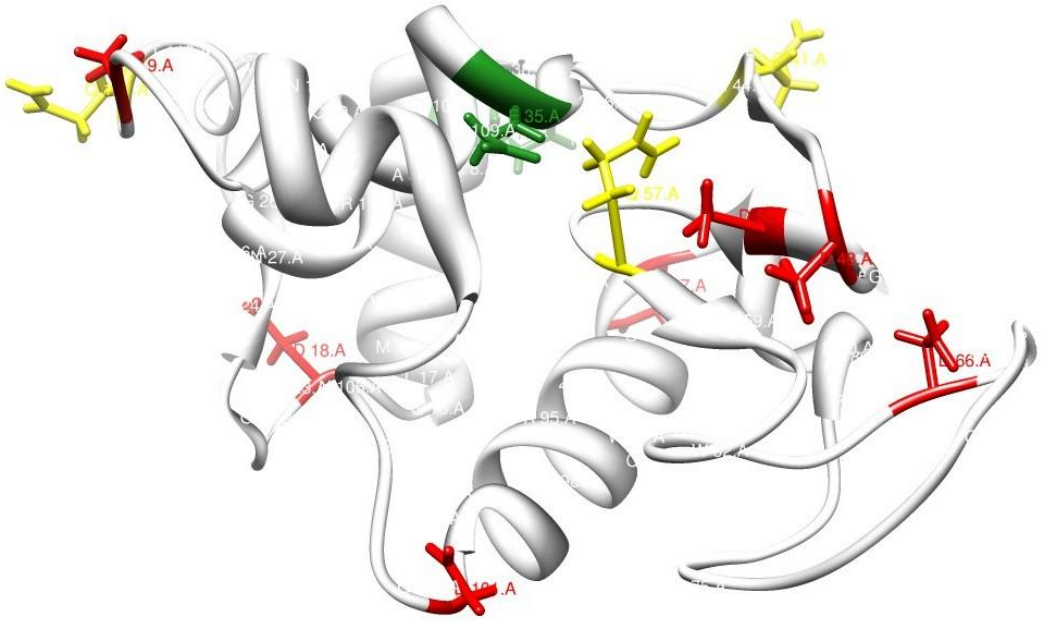


Figure 7.1 Chicken egg white lysozyme with Glu 35 (green) and Asp52 (red). In between them sits Gln 57 which is part of the b turn.

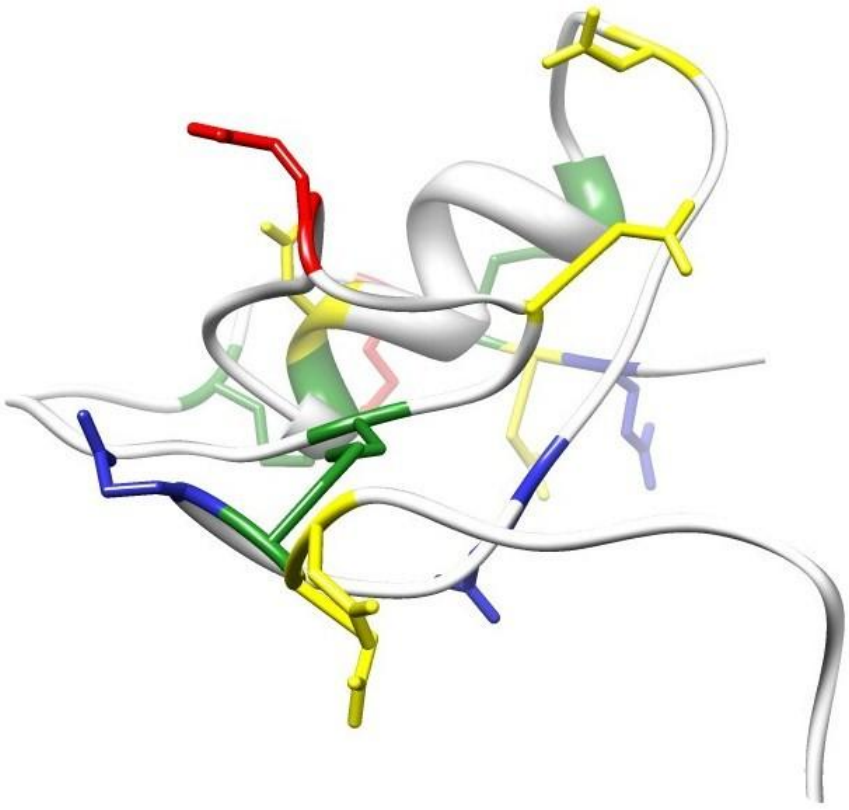


Figure 7.2 TFP4 structure with Glu51 (blue) and Asp19 (yellow), with Gln23 (red) in between them lining a pocket near the C-terminal side of TFP4

mechanism. Additionally, there are threonine residues on either side of the CXXQ hairpin which is a condition for catalysis. The structural model has to be refined in order to be able to predict which residues participate in the catalytic mechanism, but some of the conditions are already satisfied.

Serine Protease Inhibitor Function

Serine protease inhibitors diminish or abolish the activity of serine proteases. The most important classes of serine proteases are considered to be trypsin-like, elastase-like, and chymotrypsin -like enzymes so the activities of trypsin, elastase and chymotrypsin were evaluated in the presence of TFP4.

No changes in trypsin activity were observed when TFP4 was added to a trypsin-BAEE reaction. The same reaction was inhibited by the addition of known inhibitors so the conclusion was that TFP4 does not inhibit trypsin.

When TFP4 was added to an elastase- N- succinyl-(l-alanine)$_3$-p-nitroanilide, the rate of said reaction was slowed dramatically. The same result was observed when crude extract from the frontal gland secretion was incubated with elastase, so the conclusion drawn from these assays was that TFP4 is an inhibitor of elastase.

The third representative of serine proteases assayed was chymotrypsin. The enzymatic activity in the chymotrypsin-BTEE reaction was slowed down or abolished by both TFP4 and crude protein extract, revealing that TFP4 is an inhibitor of chymotrypsin.

In conclusion, the serine proteases assays provided data which strongly suggest that TFP4, the protein from the frontal gland secretion of *Coptotermes formosanus* Shiraki soldiers, is an inhibitor of elastase and chymotrypsin but does not inhibit trypsin.

Sequence alignments using BLASTp program from NCBI revealed similarities with known protease inhibitors from *Nauphoeta cinerea* and *Procambarus clarkii*. The protease inhibitor domain identified was the Kazal domain, named after L. Kazal who discovered the first pancreatic secretory trypsin inhibitory (PSTI).

Laskowski and Kato (1980) described the consensus sequence and inhibition mechanism for Kazal-type inhibitors. The consensus sequence comprises 10-12 amino acid residues which determine the inhibitor's function. P_1 is the reactive site residue of the protease inhibitor and generally corresponds to the specificity of the cognate enzyme. Inhibitors with lysine and arginine in the P1 site inhibit trypsin-like enzymes, those with tyrosine, phenylalanine leucine and methionine inhibit chymotrypsin-like enzymes and those with alanine and serine in the reactive P_1 site inhibit elastase-like enzymes (Laskowski *et al.,* 1971). Later, it was shown that inhibitors with leucine and methionine in the P1 site are strong inhibitors of elastase (Kato *et al.,* 1978).

Analysis of TFP4 sequence reveals that there is a stretch of 10 amino acid residues with strong similarity to a Kazal domain sequence from Cys4 to Cys16. The putative P_1 residue of TFP4 is methionine (Met10) which, theoretically, according to Lastkowski and Kato would give TFP4 specificity for chymotrypsin and elastase but not trypsin. This is what was actually observed in the serine protease inhibition assays performed. TFP4 inhibited both chymotrypsin and elastase but could not inhibit trypsin.

Figure 7.3 Alignment ofTFP4 and the consensus sequence for the inhibitory loop of Kazal inhibitors.

```
TFP4            CPMIYAPIC  10
                  CP IY P+C
CONSENSUS       CPRIYNPVC  10

            Score = 23.5 bits (48),  Expect = 8e-06
Identities = 6/9 (67%), Positives = 7/9 (78%), Gaps = 0/9 (0%)
```

Ozawa (1966) argues the reactive site has to be encompassed by at least 1 disulfide bridge, a structure which is present in TFP4 between cys4 and cys16. Additionally, the prolines in P_{-2} and P_4' help to ensure proper reactive site geometry (Laskowski and Kato, 1980). Both the P_{-2} and P_4' prolines are present in the TFP4 sequence.

The most common element of serine protease inhibitors' structure is the reactive site loop. This loop is exposed from the protein's globular structure and adopts a conformation which is complementary to the enzyme's surface (Bode and Huber, 1992).

In the structure presented in figure 7.4, the consensus sequence is in red with the P1 methionine in green.

This region of TFP4 is the reactive site loop described by Bode and Huber (1992) for known serine protease inhibitors, with the reactive site Met10 protruding out of the loop's plane, available to fit into the catalytic site of chymotrypsin or elastase.

Finally, in the present study data for a Lineweaver-Burke plot was recorded in order to determine the mechanism of inhibition. Reaction rates were calculated from assays performed at various concentrations of substrate and inhibitor while the enzyme concentration was constant. The resulting plot was compatible with the competitive inhibition mechanism.

The results obtained were consistent with those described previously for the inhibition mechanism for serine protease inhibitors. The mechanism resembles the hydrolysis of a singular peptide bond in regular protein substrates and is strictly a competitive inhibition mechanism for all serine proteases (Laskowski and Kato, 1980).

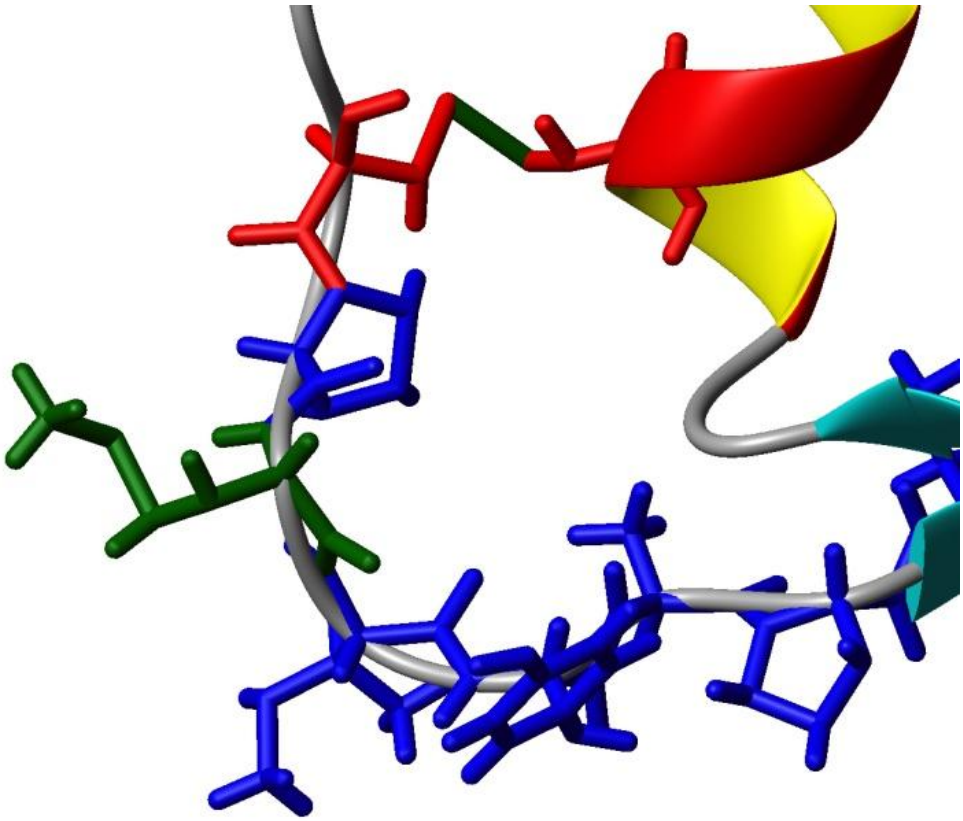


Figure 7.4 Inhibitory loop of TFP4. Methionine (Met10), pictured in green, provides inhibitory specificity for chymotrypsin and elastase. Cys4 is pictured in red forming a disulfide bond with Cys16sxc and the other loop amino acid residues are pictured in blue.

In conclusion, we strongly suggest the existence of a protein of 6853 KDa with dual functions, lysozyme and serine protease inhibitor (chymotrypsin and elastase) in the defensive secretion from the frontal gland of *Coptotermes formosanus* Shiraki soldiers. The inhibitor function is assigned to an exposed loop from Cys4 to Cys16 which functions as a substrate analog and employs a competitive inhibition mechanism for the inhibition of chymotrypsin while the lysozyme function is putatively assigned to a pocket

between the β-hairpin structure containing Gln23 and the region at the C-terminus containing Glu45.

BIBLIOGRAPHY

1. Audy P, Grenier J and Asselin A. Lysozyme activity in animal extracts after sodium dodecyl sulfate–polyacrylamide gel electrophoresis. Comp Biochem Physiol B, 92 1989, pp. 523–527

2. Balei M. Basic 1H-13C NMR spectroscopy ISBN 0444518118

3. Bieth, J, Spiess B, and Wermuth C. Thesynthesis and analytical use of a highly sensitive and convenient substrate of elastase. Biochem Med 1974;11, 350

4. Blake CC, Koenig DF, Mair GA, North AC, Phillips DC, Sarma VR. Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. Nature 206 1965; (986): 757–61

5. Blake CH. Ants preying on termites. Entomology News 1941; 52, 38 1965

6. Bode W and Huber R. Natural protein proteinase inhibitors and their interaction with proteinases. Eur J Biochem, 1992; 204: 433–451

7. Calaby JH. Observation on the banded anteater *Myrmecobius fasciatus* Waterhouse, with particular reference on its food habits. Proc Zool Soc 1960; 135: 183-207

8. Canfield RE. The Amino Acid Sequence of Egg White Lysozyme. J Biol Chem 238 1963 8: 2698–2707

9. Carpenter GD. The fly *Bengalia depressa* Walk attacking a wingless termite. Proc Entomol Soc 1919; 52-58

10. Carter P and Wells JA. Dissecting the catalytic triad of a serine protease. Nature 1998;332, 564–568

11. Deligne J, Quennedey A, Blum MS. The enemies and defense mechanisms of termites. Social Insects,1992; 2: 1-67

12. Deligne J. Mecanique de comportament de combat chez les soldats de termite. Forma Functio 1971; 4: 176-187

13. Deligne J, Quennedey A and Blum M. The enemies and defense mechanisms of termites. Social Insects, vol. 2 Academic Press, New York 1981; 1–76

14. Di Cera E. Serine Proteases Life 2009; 61(5): 510-515

15. Dimmit MA. Terrestrial ecology of spadefoot toads (*Scaphiopus*) Emergence cues, nutrition and burrowing habits. 1975 PhD Dissertation UC Riverside

16. Dracott CH. Notes on the flying white ant and scorpions that feed on them. J Bombay Nat Hist Soc 1919; 26: 873-874

17. Engel MS. and Krishna K. Family-group names for termites (Isoptera). American Museum Novitates 2004; 3432:1–9

18. Feinstein G, Kupfer A, and Sokolovsky M. N-Acetyl-(L-Ala)3-*p*-Nitroanilide as a new chromogenic substrate for elastase. Biochem Biophys Res Commun 1973; 50: 1020

19. Fleming A. On a remarkable bacteriolytic element found in tissues and secretion. Proc Roy Soc Ser B 1922; 93: 306-317

20. Ford J, Chambers R,Cohen W. An active site titration method for immobilized trypsin, Biochem Biophys Acta1973; 309:175

21. Guido AS and Ruffinelli A. Primer catalogo de los parasitos y predatores encontrado en el Uruguay. Proc Int Congr Entomol 10th, 1958; 4: 919

22. Hardt, M, Guo Y, Laine RA, Henderson G. Zymogram with Remazol brilliant blue-labeled *Micrococcus lysodeikticus* cells for the detection of lysozymes: example of a new lysozyme activity in Formosan termite defense secretions. Anal Biochem 2003; 312(1): 73-76

23. Hedstrom, L. An overview of serine proteases. Curr Protoc Protein Sci 2002; Chapter 21: Unit 21

24. Hengen P. Purification of His-Tag fusion proteins from *Escherichia coli*. Trends in Biochemical Sciences 1995; 20 (7): 285–6

25. Hill GF. On some Australian termites of the genera *Drepanotermes, Hamitermes and Leucotermes*. Bull Entomol Res 1922; 12: 363-399

26. Hojo M, Morioka M, Matsumoto T, Miura T. Insect biochemistry and molecular biology 2005; 35(4) : 347-354

27. Hummel B. A modified spectrophotometric determination of chymotrypsin, trypsin and thrombin. Can J Biochem Physiol 1959; 37:1393

28. Ikenaka T and Odani S. Studies on soybean trypsin inhibitors XI. Complete amino acid sequence of a soybean trypsin-chymotrypsin-elastase inhibitor, C-II. J Biochem 1978; 83: 207-216

29. Imler JL, Tauszig S, Jouanguy E, Forestier C, Hoffman JA. LPS-induced immune response in *Drosophila.* J Endotoxin Res 2000; 6: 459-462

30. Ishikawa Y and Aonuma H. Soldier-specific modification of the mandibular motor neurons in termites. PLoS One 2007; 3(7): e2617

31. Ito Y, Yamada H and Imoto T. Colorimetric assay for lysozyme using *Micrococcus luteus* labeled with a blue dye, Remazol brilliant blue R, as a substrate. Chem Pharm Bull 1992; 40, 1523–1526

32. Iwanaga S. The limulus clotting reaction. S Curr Opin Imunnol 1993; 5, 74-82

33. Jolles P and Jolles J. What's new in lysozyme research? Always a model system, today as yesterday. Mol Cell Biochem 1984; 63(2): 165-189.

34. Jolles B, Refregiers M, and Laigle A. Opening of the extraordinarily stable mini-hairpin d (GCGAAGC). Nucleic Acids Res. 1997; 25, 4608–4613

35. Jolles P. Lysozymes: Model enzymes in biochemistry and biology. 1996; Birkhauser, Basel

36. Kalshoven LG. Additional note on the giant elaterid *Oxynopterus mucronatus*, a predator on termites in Java. Entomol Ber 1955; 15, 273-278

37. Kato I, Kohr WJ, Laskowski M. Regulatory proteolytic enzymes and their inhibitors. Pergamon Press, New York 1978; 197-206

38. Kemp W. Organic spectroscopy (3rd edition). WH Freeman & Co., London 1991

39. Kingdom J. East African Mammals. Vol1, Academic Press, New York 1971

40. Koshland, DE. Stereochemistry and mechanism of enzymatic reactions. Biol. Rev. 1953; 28, 416–436

41. Krishna, K. Taxonomy, phylogeny, and distribution of termites, 1970;pp. 127-132, Biology of Termites, Vol. 2, Academic Press, New York.

42. La Fage JP. Practical considerations of the Formosan subterranean termite in Louisiana: a 30-year-old problem. Biology and control of the Formosan subterranean termite. Proceedings of the International Symposium on the Formosan Subterranean Termite. Hawaii Institute of Tropical Agriculture and Human Resources, 1985; pp. 37-42

43. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 1970; Aug 15;227(5259):680-5

44. Laskowski M.: Methods in Enzymology Biochem Biophys Acta 1955; 16 570.

45. Laskowski M and Kato I. Protein inhibitors of proteinases. Annu Rev Biochem 1980; 49 593–626

46. Leatherbarrow RJ. Using linear and nonlinear regression to fit biochemical data. Trends in Biochemical Sciences 1990; 15:455-458

47. Leclerc D and Asselin A. Detection of bacterial cell wall hydrolases after denaturing polyacrylamide gel electrophoresis. Can J Microbiol 1989; 35: 749–753

48. Li Y, Roy A, Zhang Y. HAAD: A Quick Algorithm for Accurate Prediction of Hydrogen Atoms in Protein Structures, PLoS One, 2009; 4: e6701

49. Little J W Autodigestion of lexA and phage lambda repressors. Proc Natl Acad Sci USA 1984 81 (5): 1375–1379

50. Lo, N. Evidence for cocladogenesis between diverse dictyopteran lineages and their intracellular endosymbionts. Molecular Biology and Evolution, 2003; 20, 907–913

51. *Lo* N, *Hayashi* Y, Kitade O. *Should environmental caste* determination be assumed for termites? Am Nat 2009 Jun;173(6):848-53

52. Maeterlinck, M. The Life of the White Ant. Dodd, Mead, New York 1939

53. Miura T, Kamikouchi A, Sawata M, Takeuchi H, Natori S, Kubo T, Matsumoto T: Soldier caste-specific gene expression in the mandibular glands of *Hodotermopsis japonica*. *Proc Natl Acad Sci USA* 1999, 96:13874-13879

54. Mao L, Henderson G, Liu Y, Laine RA. Formosan subterranean termite (Isoptera: Rhinotermitidae) soldiers regulate juvenile hormone levels and caste differentiation in workers. Ann Entomol Soc 2005; 98: 340–345

55. Mill AE. Foraging and defensive behaviour in neotropical termites. PhD thesis. 1982; Univ. of Southampton, England.

56. Moore BP. Pheromones in termite societies, American Elsevier, New York 1974; 250-265

57. Nakata, N., Tobe, T., Fukuda, I., Suzuki, T., Komatsu, K., Yoshikawa, M., Sasakawa, The absence of a surface protease, OmpT, determines the intercellular spreading ability of Shigella: the relationship between the ompT and kcpA loci. Mol Microbiol 1993 Aug; 9(3):459-68

58. NCBI BLAST [http://www.ncbi.nlm.nih.gov/BLAST].

59. Noirot C. Caste differentiation in Isoptera: basic features, roles of pheromones. Ethol Ecol Evolut.1991; Sp Iss 1:3-7

60. Noirot C, Quennedey A. Glands, gland cells, glandular units: Some comments on terminology and classification. Annu Soc Entomol Fr 1991; 27: 123–128.

61. Osbrink W, Woodson, WD, Lax AR. Populations of Formosan subterranean termite, Coptotermes formosanus (Isoptera: Rhinotermitidae), established in living urban trees in New Orleans, Louisiana, U. S. A., 3rd International Conference on Urban Pests, Czech Republic, Graficke zavody Hronov, 1999; 341-345.

62. Osserman EF and Lawlor DP. Serum and urinary lysozyme (muramidase) in monocytic and monomyelocytic leukemia. J. Exp. Med., 1966; 124: 921–952

63. Prestwich GD. Termites: Dwellers in the dark. Natl Geographic. 1978; 153: 532-547

64. Prestwich GD. Chemical defense by termite soldiers. Journal of Chemical Ecology 1979; (5) 3: 459-480

65. Prestwich, GD. Defense mechanisms of termites. Annu Rev Entomol 1984; 29: 201–232

66. Potvin C, Leclerc D, Tremblay G, Asselin A. Bellemare G. Cloning, sequencing and expression of a Bacillus bacteriolytic enzyme in *Escherichia coli*. Mol Gen Genet 1988; 214: 241–248

67. Puyet A, Greenberg B, Lacks SA. Genetic and structural characterization of endA. A membrane-bound nuclease required for transformation of *Streptococcus pneumoniae*. J Mol Biol 1990 Jun 20;213(4):727–738

68. Quennedey A. Observations cytologique et chimique sur la glande frontale des termites. Proc. 7th Congr. lnt. Union Study Soc Insects, London, 1973; 324-26

69. Quennedey A. The labrum of *Schedorhinotermes minor* soldier: morphology, innervation, and structure. Cell Tissue Res 1975;1 60:8 1-98

70. Quennedey A. Morphology and ultrastructure of termite defense glands. Defensive Mechanisms of Social Insects 1983; 27: 123–128

71. Quennedey A. Morphology and ultrastructure of termite defense glands, Defensive mechanisms in social insects, Praeger Publishers, New York (1984), 151–200

72. Page MJ and DiCera E. Serine peptidases: classification, structure and function. Cell Mol Life Sci 2008; 65: 1220-1236

73. Rawlings ND, Morton FR. MEROPS: The Protease Database. Nucleic Acids Research 2008; 36 Database issue, D320–325

74. Rigby M. Procedings of the International Research Conference on Proteinase Inhibitors, de Gruyter, Berlin 1971; 74-88

75. Rinderknecht H, Widding P and Haverback B, A new method for the determination of α-amylase. Experientia 1967; 23: 805

76. Rowan MK. The foods of South African birds. Ostrich, Suppl 1971; 8, 343-356

77. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols 2010 5: 725-738

78. Scharf ME, Wu-Scharf B, Pittendrigh R, Bennett GW. Caste- and development-associated gene expression in a lower termite. Genome Biol., 4 2003; R62

79. Scharf ME, Buckspan CE, Grzymala TL, Zhou X Regulation of polyphenic caste differentiation in the termite *Reticulitermes flavipes* by interaction of intrinsic and extrinsic factors. J Exp Biol 2007; 210:4390–4398

80. Scharf ME, Kovaleva ES, Jadhao S, Campbell JH, Buchman GW. Functional and translational analyses of a beta-glucosidase gene (glycosyl hydrolase family 1) isolated from the gut of the lower termite Reticulitermes flavipes. Insect Biochem Mol Biol 2010; 40:611-620

81. Schoofs L, Clynen E, Salzet M. Trypsin and chymotrypsin inhibitors in insects and gut leeches. Curr Pharm Des 2002 ;8 (7):483-91

82. Sheppe W. Invertebrate predation on termites of the African savanna. Insectes Soc 1970;17:205-18

83. Shugar D. The measurement of lysozyme activity and the ultra-violet inactivation of lysozyme. Biochim Biophys Acta 1952; 8: 302–309

84. Silverstein A, Bassler B, Morrill C. Spectrophotometric identification of organic compounds (4^{th} edition). ISBN 047109700

85. Spaton SG and Prestwich GD. Chemical defense and self-defense: biochemical transformations of contact insecticides produced by soldier termites. Tetrahedron. 1981; 38: 1921–1930

86. Stamm OA, Zur reaktion von reaktivfarbstoffen mit Cellulose: II. Natur der binding. Helv Chim Acta, 1963; 46: 3008

87. Stein R and Trainor D. Mechanism of Inactivation of Human Leukocyte Elastase by a Chloromethyl Ketone: Kinetic and Solvent Isotope Effect Studies , Biochem 1986;25, 5414

88. Stewart K. A method for automated analyses of the activities of trypsin, chymotrypsin and their inhibitors. Anal Biochem 1973; 51: 11

89. Stuart AM. The role of chemicals in termite communication, Advances in Chemoreception, Appleton-Century-Crofts, New York 1970; 2: 79-106

90. Tarver MR, Schmelz EA Effects of soldier-derived terpenes on soldier caste differentiation in the termite Reticulitermes flavipes. J Chem Ecol 2009; 35(2): 256-264

91. Tian M, Huitema E, Da Cunha L, Torto-Alalibo T, Kamoun SA. Kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B. J Biol Chem 2004;279: 26370–26377

92. Vargo E and Husseneder C. Biology of Subterranean Termites: Insights from Molecular Studies of *Reticulitermes* and *Coptotermes* Annual Review of Entomology 2009; 54: 379-403

93. Vocadlo DJ, Davies GJ , Laine RA, Withers SG. Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate Nature 2001;412, 835-838

94. Walsh DA, Ashby CD, Gonzalez C, Calkins O, Fischer EH, Krebs EG. Purification and characterization of a protein inhibitor of adenosine 3',5'-monophosphate-dependent protein kinases. J Biol Chem 1971; 246: 1977-1985

95. Wang Q, Grahal RW, Trimbur D, Warren RAJ, Withers SG. Changing enzymatic reaction mechanisms by mutagenesis – conversion of a retaining glucosidase into an inverting enzyme. J Am Chem Soc 1994; 116: 11594–11595

96. Ware, JL. Relationships among the major lineages of *Dictyoptera*: the effect of outgroup selection on dictyopteran tree topology. Systematic Entomology, 2008;33, 429–450

97. Weber NA. Termite prey of some African ants. Entomol News 1964; 74:197-204

98. Weesner FM. Evolution biology of termites. Annual Review of Entomology. 1960;5; 153-170

99. Wilcox PE. Chymotrypsinogens — chymotrypsins. Methods in Enzymology 1970;19: 64–108

100. Wilson EO. The Insect Societies. Belknap Press, Cambridge 1971

101. Wohlkönig A, Huet J, Looze Y, Wintjens R, Structural relationships in the lysozyme superfamily: significant evidence for glycoside hydrolase signature motifs. PLoS ONE 2010;5(11): e15388

102. Wood TG. The effects of clearing and grazing on the termite fauna (Isoptera) of tropical savannas and woodlands. Progress in Soil Zoology, Procs 5 Intl Coll on soil zoology 1975; 409-413

103. Wood WF, Truckenbrodt W, Meinwald J. Chemistry of the defensive secretion from the African termite *Odontotermes badius*. Ann Entomol Soc Am 1975; 68:359-360.

104. Wu S,and Zhang Y. LOMETS: A local meta-threading-server for protein structure prediction. Nucleic Acids Research 2007; 35: 3375-3382

105. Wu S and Zhang Y. MUSTER: Improving protein sequence profile-profile alignments by using multiple sources of structure information. Proteins 2008; 72: 547-556

106. Xu D andZhang Y. Generating Triangulated Macromolecular Surfaces by Euclidean Distance Transform. PLoS ONE2009; 2009; 4: e8140

107. Zhang Y and Skolnick J. SPICKER: Approach to clustering protein structures for near-native model selection, Journal of Computational Chemistry 2004; 25: 865-871

108. Zhang Y, NW-align http://zhanglab.ccmb.med.umich.edu/NW-align

109. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 2008; 9, 40

110. Zhou X, Tarver MR, Scharf ME. Hexamerin-based regulation of juvenile hormone-dependent gene expression underlies phenotypic plasticity in a social insect. Development 2007; 134:601–610

111. Zhou X and Tarver MR. Two hexamerin genes from the termite *Reticulitermes flavipes*: Sequence, expression, and proposed functions in caste regulation. Gene 2006;376(1): 47-58

112. Zubay GL. Biochemistry. Addison-Wesley Publishing Company Inc. 1983; 143-158

VITA

Ioan Horia Negulescu was born in Braila, Romania. He graduated from Liceul National in Iasi, Romania, with a Bacalaureat (High School Diploma) in June 1987. Horia then attended “Gr. T. Popa” Medical School in Iasi, Romania, from where he transferred to Case Western Reserve University in Cleveland, Ohio, and graduated in May 1994 with a bachelor of arts degree in biology with a minor degree in French literature. He worked as a research associate in the Molecular Biology Department at Princeton University in Princeton, New Jersey, and, in the fall of 2004, he entered the Biochemistry and Molecular Biology Program at Louisiana State University, working toward the doctor of philosophy degree under the direction of Dr. Roger A. Laine.